

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 12:59:49 ; Search time 9150.23 Seconds
(without alignments)
10942.062 Million cell updates/sec

Title: US-09-303-232-1_COPY_372_2681
Perfect score: 2310
Sequence: 1 atgaaaatgcacactgaa.....caccacatattattgtctcg 2310

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_btg.*
- 3: gb_in.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2310	100.0	2886	6	E58346 Nucleic aci
2	2310	100.0	2886	6	AX009610 Sequence
3	2187.4	94.7	2907	3	AF272778 Drosophil
4	1994.8	86.4	2834	3	AY036613 Drosophil
5	776.2	33.6	42079	2	AC017620 Drosophil
6	776.2	33.6	158758	3	AC092243 Drosophil
7	776.2	33.6	186803	3	AC092225 Drosophil
8	776.2	33.6	272521	3	AE003642 Drosophil
9	776.2	33.6	320754	3	DROSADH02
10	515.4	22.3	1683	3	DMES54210
11	512.8	22.2	3629	3	AF143846
12	512.8	22.2	3700	6	AX009612 Sequence
13	512.8	22.2	3701	6	E58347 Nucleic aci
14	448	19.4	2023	3	AF321445 Drosophil
15	441.6	19.1	2023	3	AF321447 Drosophil
16	436.8	18.9	2023	3	AF321446 Drosophil
17	433.6	18.8	1665	3	DME554209
18	433.6	18.8	1699	3	BT011147 Drosophil
19	424.6	18.4	2110	3	AF321449 Drosophil
20	411.4	17.8	3029	3	AF143847 Heliothis
21	411.4	17.8	3109	6	E58348 Nucleic aci
22	411.4	17.8	3109	6	AX009614 Sequence
23	389.8	16.9	2068	3	AF321448 Drosophil
24	321	13.9	1461	5	AY299465 Takifugu
25	320.2	13.9	2116	3	AY036614 Drosophil
26	316.4	13.7	1542	5	AY299466 Takifugu
27	314	13.6	1719	5	AY247962 Dario rer
28	313.8	13.6	2106	10	RATNARAD
29	313.8	13.6	3030	10	S53987
30	311.4	13.5	1668	5	AY298752
31	307.4	13.3	1848	10	MUSNARS
32	305.8	13.2	2088	10	AF225980
33	297.8	12.9	1559	9	HUMA7NAR
34	297.8	12.9	1916	9	AF385585
35	297.8	12.9	1977	9	HS040583
36	297.8	12.9	2087	9	HSARA7A
37	296.2	12.8	1509	6	AX054567
38	296.2	12.8	1509	9	HSNACHRA7
39	296.2	12.8	1555	9	HSCHRNA7A
40	296.2	12.8	1665	5	AY298753
41	296.2	12.8	1876	6	AR055255
42	296.2	12.8	1876	6	AR071403
43	296.2	12.8	1876	6	AR173187
44	296.2	12.8	1876	6	AR224030
45	296.2	12.8	1876	6	AR282833

ALIGNMENTS

RESULT 1
E58346
LOCUS E58346
DEFINITION Nucleic acid encoding insect actyl choline receptor subunit.
ACCESSION E58346
VERSION E58346.1 GI:13019345
KEYWORDS JP 2000023680-A/1.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2886)
AUTHORS Martin,A., Nadja,E. and Thomas,S.
TITLE Nucleic acid encoding insect actyl choline receptor subunit

2886 bp DNA linear PAT 18-JUN-2001


```
Db 2172 GGACGACGCTGATCTCTAGAGTCCGACACACGCCCTGTTGCGACACATCTCCGACGG 2231
Qy 1861 AAGCACCAGATATCTCCGACGTTGAGTGAAGAGAGCGCTCGTGAATCGCTCTGGCC 1920
Db 2232 AAGCACCAGATATCTCCGACGTTGAGTGAAGAGAGCGCTCGTGAATCGCTCTGGCC 2291
Qy 1921 AACGTACTAGACATGATGATCTTCGGGACAAATTTGTCGCCCATGAGCCCGCGGA 1980
Db 2292 AACGTACTAGACATGATGATCTTCGGGACAAATTTGTCGCCCATGAGCCCGCGGA 2351
Qy 1981 ACATGCGCACACACCCCGCTTCTATCGACGCTTTATGACAGAGCGAGCATGGCAGC 2040
Db 2352 ACATGCGCACACACCCCGCTTCTATCGACGCTTTATGACAGAGCGAGCATGGCAGC 2411
Qy 2041 ATTGGGCAATTTGCGACACCCGAAATCGCGATCGCGTACCCCATCATAGCTGATCAAA 2100
Db 2412 ATTGGGCAATTTGCGACACCCGAAATCGCGATCGCGTACCCCATCATAGCTGATCAAA 2471
Qy 2101 TCATCACTGAATGATGATGATGATCTTAAAGGAAATTCGTTTATTAACATGATCAG 2160
Db 2472 TCATCACTGAATGATGATGATGATCTTAAAGGAAATTCGTTTATTAACATGATCAG 2531
Qy 2161 CTAGTAAAGATGACGAGTGAATGACATTCGCAATGATGGAATTTGCGAGCTATGGTC 2220
Db 2532 CTAGTAAAGATGACGAGTGAATGACATTCGCAATGATGGAATTTGCGAGCTATGGTC 2591
Qy 2221 GTTGACAGATGTCCTTATCATATTCACAAATGTCGAATATTTAGCCACAAATAGCTGTA 2280
Db 2592 GTTGACAGATGTCCTTATCATATTCACAAATGTCGAATATTTAGCCACAAATAGCTGTA 2651
Qy 2281 CTACTATCGGCACACATATTTATGTCG 2310
Db 2652 CTACTATCGGCACACATATTTATGTCG 2681
```

```
RESULT 2
AX009610
LOCUS AX009610 2886 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent EP0962528.
ACCESSION AX009610
VERSION AX009610.1 GI:9996842
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Adamczewski M.D., Schulte T.D. and Oellers N.D.
TITLE Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL Patent: EP 0962528-A 1 08-DEC-1999;
BAYER AG (DE)
FEATURES
source Location/Qualifiers
1..2886
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"
372..2684
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:9996843"
/db_xref="FEMTREMEL:CA07499"
/translation="MKNAQLKTEVDDDLWLAIVLHCSNFSSSSSTRITSSNQHR
NQQLTIQPSRLSTPKHSNIASEQHSQQEPKASDEVDVHGRSNDQOHLQQLDSS
NMLSPKTAARATDATTQCPNIRLCARORLRNRKRKPATNETDICKQQL
SMPPKTKSTDTYSTPATTCPTATYMOCRASDNESPIPSHEDRVSTATFANVLI
VLQVLLVSLQWQLHVQORSVLLFRRIAATFIYSIGSFAQLKNSSSSSNS
NDSQTILINGHWSIIFLLYLNLISAKVLGAEKRLHLHLLDLPYNTLPERVLNES
DQLQSLGLTLMQIIDVENKQLLVNWLKLEWMDNLRWNTSDYGVKDLRIPPHR
IWKPDVLMYASADEGFDGTQTNVVRNNGSLYVPPGIFKSTCKIDITWFPFDQRC
EMKFGSWTYDGFQDLQLQDETGGDISSVLNGEWELLGVFGKNEIYYNCCPEPYID
```

CDS

```
ITFAIIRRTLYFFNLIIPCVLLIASMALLGFTLPPDSGEKLSLGVITILLSLTVFLN
MVATETPATSDAVPLMIRIVFLCWLFWILNRSRFRLEILIEFPTTCCSDTSSEKHKQI
LSVEIKRSKSSLLANLIDDDFRHNCRPMTPGGTLPHNPAPFYRTVYGGDGSIG
PIGSTMPDAVTHHTTCIKSSTYELGLILKEIRFITDQLRKDDCEANDIANDWKEAAMV
VDRCLLIIFTMFAILATIIVLLSAPHIIVS"
ORIGIN
Query Match 100.0%; Score 2310; DB 6; Length 2886;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAATGCACAACTGAAACTGACTGAAGTTGACGATGATGAGCTGTGCTGCGAGTA 60
Db 372 ATGAAAAATGCACAACTGAAACTGACTGAAGTTGACGATGATGAGCTGTGCTGCGAGTA 431
Qy 61 AGATTAGCGCACTGCGACGACCACTTTAGCAGCAGTAGCAGCAACAAGAACCCACGAGC 120
Db 432 AGATTAGCGCACTGCGACGACCACTTTAGCAGCAGTAGCAGCAACAAGAACCCACGAGC 491
Qy 121 AACGAGAGGCACACACGAGCACTCAACAACCTGCAACCTGCAACCTGCAACCTGCAAC 180
Db 492 AACGAGAGGCACACACGAGCACTCAACAACCTGCAACCTGCAACCTGCAACCTGCAAC 551
Qy 181 CACAGCAACATTTGCAAGCGGACGACCAACAATAGCCAGCAACAAGGAGCCAGCATCGAAG 240
Db 552 CACAGCAACATTTGCAAGCGGACGACCAACAATAGCCAGCAACAAGGAGCCAGCATCGAAG 611
Qy 241 GAGGATGAGCAACACGAGTAGAAGCAATGACAGCAGCGCATCTGCAACAGCTAGAC 300
Db 612 GAGGATGAGCAACACGAGTAGAAGCAATGACAGCAGCGCATCTGCAACAGCTAGAC 671
Qy 301 AGCAGCAACATTTGTCGCCAAGACAGCGCGCAGCACTGCTGCGCGGATGAGCA 360
Db 672 AGCAGCAACATTTGTCGCCAAGACAGCGCGCAGCACTGCTGCGCGGATGAGCA 731
Qy 361 ACAACCCCAACAACAATAAGACTGTGTGTCAGCGCAAGCGCAACAACGATTCGCTGCG 420
Db 732 ACAACCCCAACAACAATAAGACTGTGTGTCAGCGCAAGCGCAACAACGATTCGCTGCG 791
Qy 421 CGACGAAAAAGAAAAACAGCAACCCCAACCGAAACAGATATCAAGAAACCAACAGCACTT 480
Db 792 CGACGAAAAAGAAAAACAGCAACCCCAACCGAAACAGATATCAAGAAACCAACAGCACTT 851
Qy 481 AGCATGCTGCTTCAAAAACGCGCAATCAACGAGCACTTACAGCAACAGCAAGCAACA 540
Db 852 AGCATGCTGCTTCAAAAACGCGCAATCAACGAGCACTTACAGCAACAGCAAGCAACA 911
Qy 541 ACCAGCTGTGCGACAGCACTTACATGCAATGTGCGAGCGAGCAATGAGTTTCAATTT 600
Db 912 ACCAGCTGTGCGACAGCACTTACATGCAATGTGCGAGCGAGCAATGAGTTTCAATTT 971
Qy 601 CCGATATCGAGACATGATAGATATCCACGCCACATTCGCTGGGTGTTGCAATGCTG 660
Db 972 CCGATATCGAGACATGATAGATATCCACGCCACATTCGCTGGGTGTTGCAATGCTG 1031
Qy 661 CAGTGTGCTGCTGCTGCTGCAACAGTGCACCTTCAACGCGCAACAGCACTTCACTT 720
Db 1032 CAGTGTGCTGCTGCTGCTGCAACAGTGCACCTTCAACGCGCAACAGCACTTCACTT 1091
Qy 721 CTGTTTCAAGAGGATGCGCAGCGAGCACCCTGCTTCACTTCTTCTATTTAGGCGAGCTT 780
Db 1092 CTGTTTCAAGAGGATGCGCAGCGAGCACCCTGCTTCACTTCTTCTATTTAGGCGAGCTT 1151
Qy 781 GCGCACTGAAAAATAGCAGCAGCAGTAGCAGCAGCAACAGCAGCAACAACAGCAGC 840
Db 1152 GCGCACTGAAAAATAGCAGCAGCAGTAGCAGCAGCAACAGCAGCAACAACAGCAGC 1211
Qy 841 ACCGAAATATTAACGCACTTAATAACACTCATGATATTTTATTGATATATTTGAAT 900
Db 1212 ACCGAAATATTAACGCACTTAATAACACTCATGATATTTTATTGATATATTTGAAT 1271
Qy 901 TTATCTCTAAAGTTTCTCTAGCAGGATATCATGAAAGAGACTGTTTACAGCATCTTTTG 960
```

Db 1272 TTATCTGCTAAAGTTTGCTAGCAGGATATCATGAAAGAGAGACTGTTACACGATCTTTTG 1331
QY 961 GATCCTTTATATACACTAGAACGTCCTCGTTCTCAATGAATCGAGCCCGTTACAAATTAAGC 1020
Db 1332 GATCCTTTATATACACTAGAACGTCCTCGTTCTCAATGAATCGAGCCCGTTACAAATTAAGC 1391
QY 1021 TTTGGTTTAACTTTAATGCAAAATATCGATGTGGACGAGAAAAATCAATTTGCTAGTCACT 1080
Db 1392 TTTGGTTTAACTTTAATGCAAAATATCGATGTGGACGAGAAAAATCAATTTGCTAGTCACT 1451
QY 1081 AATGTGTGGTTAAACTGGAGTGGAAACGACATGAATCTCCGCTGGAAACACCTCCCACTAT 1140
Db 1452 AATGTGTGGTTAAACTGGAGTGGAAACGACATGAATCTCCGCTGGAAACACCTCCCACTAT 1511
QY 1141 GSCGGAGTTAGGATCTCGGATACGGCGGATCGCATCTGGAAGCCGAGCGTGTGATG 1200
Db 1512 GSCGGAGTTAGGATCTCGGATACGGCGGATCGCATCTGGAAGCCGAGCGTGTGATG 1571
QY 1201 TACAAACAGTGGCGATGAGGGATTGACCGGCACCTACACAGAACGTTGGTGTGGCGAAC 1260
Db 1572 TACAAACAGTGGCGATGAGGGATTGACCGGCACCTACACAGAACGTTGGTGTGGCGAAC 1631
QY 1261 AACGGCTCGTGTCTATAGTTCCCGCCGCGGATCTTCAAGTCGACGTGCAAGATCGACATC 1320
Db 1632 AACGGCTCGTGTCTATAGTTCCCGCCGCGGATCTTCAAGTCGACGTGCAAGATCGACATC 1691
QY 1321 ACGTGTTCCTTCGATGACGAGCGGTGCGAGATGAAGTTCGGCAGTTGGACCTACGAC 1380
Db 1692 ACGTGTTCCTTCGATGACGAGCGGTGCGAGATGAAGTTCGGCAGTTGGACCTACGAC 1751
QY 1381 GGATTCACAGCTGGATTTACAATTACAAGATGAAACTGGCGGTGATATCAGCAGTTTACGTG 1440
Db 1752 GGATTCACAGCTGGATTTACAATTACAAGATGAAACTGGCGGTGATATCAGCAGTTTACGTG 1811
QY 1441 CTCACCGCGAGTGGGAACACTCGGGTGTGCCGGGAAACGTAACGAGATCTATTACAAAC 1500
Db 1812 CTCACCGCGAGTGGGAACACTCGGGTGTGCCGGGAAACGTAACGAGATCTATTACAAAC 1871
QY 1501 TGCTCCCGGACCTTATAGACATCACTTCGCCATCATCATCGCGGACGAGACACTG 1560
Db 1872 TGCTCCCGGACCTTATAGACATCACTTCGCCATCATCATCGCGGACGAGACACTG 1931
QY 1561 TACTATTTCTTCAACCTGATCATACCTTTGTGTACTGATTTGCCCTCCATGSCCTTGTCTCGA 1620
Db 1932 TACTATTTCTTCAACCTGATCATACCTTTGTGTACTGATTTGCCCTCCATGSCCTTGTCTCGA 1991
QY 1621 TTCACCTCGCCGACAGATTCGGGTGAAAAATTTATCGTGGGTGTTACCATCTTGTCTCTCG 1680
Db 1992 TTCACCTCGCCGACAGATTCGGGTGAAAAATTTATCGTGGGTGTTACCATCTTGTCTCTCG 2051
QY 1681 CTGACCGTGTCTCTGAATATGTTCCGAGACAATGCGCGCTACTTCCGATCGCGTGCCA 1740
Db 2052 CTGACCGTGTCTCTGAATATGTTCCGAGACAATGCGCGCTACTTCCGATCGCGTGCCA 2111
QY 1741 TTGTGGATACGATCGTGTGTTTTGTGCTGGTGCATGATATTCGAAATGAGTCCGCCCA 1800
Db 2112 TTGTGGATACGATCGTGTGTTTTGTGCTGGTGCATGATATTCGAAATGAGTCCGCCCA 2171
QY 1801 GGACGACCGCTGATCCTAGAGTTCCGACGACGAGCGCTGTTCCGACACATCTCCGAGCGG 1860
Db 2172 GGACGACCGCTGATCCTAGAGTTCCGACGACGAGCGCTGTTCCGACACATCTCCGAGCGG 2231
QY 1861 AAGCACGAGATCTCTCCGACGTTGAGCTGAAAGAGCGCTCGTCCGAAATCGCTGCTGGCC 1920
Db 2232 AAGCACGAGATCTCTCCGACGTTGAGCTGAAAGAGCGCTCGTCCGAAATCGCTGCTGGCC 2291
QY 1921 AACGTACTAGACATCGATGATGACTTCCGGACAAATTTGTGCGCCCATGACGCGCCGCGGA 1980
Db 2292 AACGTACTAGACATCGATGATGACTTCCGGACAAATTTGTGCGCCCATGACGCGCCGCGGA 2351
QY 1981 ACACGTGCCACAAACCCGCTTTCTATCGACGGTTTATGACAGGCGAGCATGGCAGC 2040
Db 2352 ACACGTGCCACAAACCCGCTTTCTATCGACGGTTTATGACAGGCGAGCATGGCAGC 2411

QY 2041 ATTGGGCCAAATGGCAGCACCCGAATGCGGATGCGGTCACCCATCATACGTGCATCAA 2100
Db 2412 ATTGGGCCAAATGGCAGCACCCGAATGCGGATGCGGTCACCCATCATACGTGCATCAA 2471
QY 2101 TCATCAACTGAATATGAATTAGTTTAATCTTTAAAGGAAATTCGCTTTTATACTGATCAG 2160
Db 2472 TCATCAACTGAATATGAATTAGTTTAATCTTTAAAGGAAATTCGCTTTTATACTGATCAG 2531
QY 2161 CTACGTAAGATGACGAGTGAATGACATGTCGAATGATGGAAATTTGCGAGCTATGGTC 2220
Db 2532 CTACGTAAGATGACGAGTGAATGACATGTCGAATGATGGAAATTTGCGAGCTATGGTC 2591
QY 2221 GTTGACAGACTGTGCTTTATCATATTTCAATGTTTCAATATTTAGCCACCAATAGCTGTA 2280
Db 2592 GTTGACAGACTGTGCTTTATCATATTTCAATGTTTCAATATTTAGCCACCAATAGCTGTA 2651
QY 2281 CTACTATCGGCACCAATATTTATTTCTCG 2310
Db 2652 CTACTATCGGCACCAATATTTATTTCTCG 2681

RESULT 3

AF272778

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 2907

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/chromosome="2"

/map="2L; 34B4-5"

/dev_stage="embryo"

1. 2907

/gene="nAcRalpha-34E"

/allele="A"

289. .2712

/gene="nAcRalpha-34E"

/note="ion channel; neurotransmitter transmembrane receptor"

/codon_start=1

/product="nicotinic acetylcholine receptor Dalphas subunit"

/protein_id="AA013390.1"

/db_xref="GI:20152840"

/translation="MKNAQLKLTVEVDDDELMLAVLHLCSSNFSSSSSTRITSSNQRH NQQLITLQPSRLSTKHENLASEQNSQQSPASKEDVANHGRSNDQDTHLQDLSS NMLSPKTAATAAGDEATQPTIRLCARRKRLRRKKRPATPNETDILKKQQL SMPFPKTKXSDTSTYTPAATTCPTATYMQCRASDNFSPISIRHDRVSTATFAWLH

AF272778 2907 bp mRNA linear INV 29-APR-2002
Drosophila melanogaster nicotinic acetylcholine receptor Dalphas
subunit (nAcRalpha-34E) mRNA, nAcRalpha-34E-A allele, complete cds.
AF272778
AF272778.1 GI:20152839

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 2907)

Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.

Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,

Dalphas, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a

New and Highly Conserved Target of Adenosine Deaminase Acting on

RNA-Mediated A-to-I Pre-mRNA Editing

Genetics 160 (4), 1519-1533 (2002)

11973307

21969411

2 (bases 1 to 2907)

Grauso, M. and Sattelle, D.B.

Direct Submission

Submitted (27-MAY-2000) Human Anatomy and Genetics, MRC-FGU,

University of Oxford, South Park Road, Oxford OX1 3QX, UK

Location/Qualifiers

1. 2907

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/chromosome="2"

/map="2L; 34B4-5"

/dev_stage="embryo"

1. 2907

/gene="nAcRalpha-34E"

/allele="A"

289. .2712

/gene="nAcRalpha-34E"

/note="ion channel; neurotransmitter transmembrane receptor"

/codon_start=1

/product="nicotinic acetylcholine receptor Dalphas subunit"

/protein_id="AA013390.1"

/db_xref="GI:20152840"

/translation="MKNAQLKLTVEVDDDELMLAVLHLCSSNFSSSSSTRITSSNQRH

NQQLITLQPSRLSTKHENLASEQNSQQSPASKEDVANHGRSNDQDTHLQDLSS

NMLSPKTAATAAGDEATQPTIRLCARRKRLRRKKRPATPNETDILKKQQL

SMPFPKTKXSDTSTYTPAATTCPTATYMQCRASDNFSPISIRHDRVSTATFAWLH

QY 181 CACAGCAACATTGCAAGCGAGCAGCACATATAGCCACAGCAGGAGCCAGCATCGAAGAC 240
DB 469 CACAGCAACATTGCAAGCGAGCAGCACATATAGCCACAGCAGGAGCCAGCATCGAAGAC 528
QY 241 GAGGATGTAGCCAAACCCAGGTAGAGCAATGACAGCAGCAGCATCTGCAACAGCTAGAC 300
DB 529 GAGGATGTAGCCAAACCCAGGTAGAGCAATGACAGCAGCAGCATCTGCAACAGCTAGAC 588
QY 301 AGCAGCAACATGTTGTCGCAAGAGCAGCGCAGCAGCAACTGTGTCGCGGAGATGAAGCA 360
DB 589 AGCAGCAACATGTTGTCGCAAGAGCAGCGCAGCAGCAACTGTGTCGCGGAGATGAAGCA 648
QY 361 CAAACCCCAACCAACCAACCAATATAGACTGTGTCAGCAGCAGCAGCAAGTGTGCTGC 420
DB 649 CAAACCCCAACCAACCAACCAATATAGACTGTGTCAGCAGCAGCAGCAAGTGTGCTGC 708
QY 421 CGACGAAAGAAACCCAGCAACCCCAACCAAGAAACAGATATCAAGAAACCAACAGCAACTT 480
DB 709 CGACGAAAGAAACCCAGCAACCCCAACCAAGAAACAGATATCAAGAAACCAACAGCAACTT 768
QY 481 AGCATGCTCCCTTCAAAACCGGCAAAATCCAGGACACCTCAGCAGCAGCAGCAACA 540
DB 769 AGCATGCTCCCTTCAAAACCGGCAAAATCCAGGACACCTCAGCAGCAGCAGCAACA 828
QY 541 ACCAGCTGTCGACAGCCACCTACATGCAATGTGAGCGCAGCAGCAATGAGTTCAGTATT 600
DB 829 ACCAGCTGTCGACAGCCACCTACATGCAATGTGAGCGCAGCAGCAATGAGTTCAGTATT 888
QY 601 CCGATATCGAGACATGATAGAGTATCCAGGCGCACATTCGGCTGGGTGTGTCATGTGCTG 660
DB 889 CCGATATCGAGACATGATAGAGTATCCAGGCGCACATTCGGCTGGGTGTGTCATGTGCTG 948
QY 661 CAGGTGCTGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 949 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 721 CTGTTTCAAGAGATCGCAGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1008
DB 1009 CTGTTTCAAGAGATCGCAGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1068
QY 781 GCGCAACTGAAAAATAGCAGC --- AGCAGCAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGC 837
DB 1069 GCGCAACTGAGGAGCAGCAGCAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1128
QY 838 AGCAGCAAAATATTAACGGACTTAATTAACACTCATGAGATATTTTATGATATATTTG 897
DB 1129 AGCAGCAAAATATTAACGGACTTAATTAACACTCATGAGATATTTTATGATATATTTG 1188
QY 898 AATTTATCTGCTAAAGTTTGCTAGCAGGATATCATGAAAGAGACTGTACAGCATCTT 957
DB 1189 AATTTATCTGCTAAAGTTTGCTAGCAGGATATCATGAAAGAGACTGTACAGCATCTT 1248
QY 958 TTGATCCTTATATATACATAGAACGTCCTGCTTCTCAATGAATCGGACCCGCTTACAATTA 1017
DB 1249 TTGATCCTTATATATACATAGAACGTCCTGCTTCTCAATGAATCGGACCCGCTTACAATTA 1308
QY 1018 AGCTTTGTTTAACTTTAATGCAAAATTCATGATGAGCAGGAAATTCATTTGCTAGTC 1077
DB 1309 AGCTTTGTTTAACTTTAATGCAAAATTCATGATGAGCAGGAAATTCATTTGCTAGTC 1368
QY 1078 ACTAATGTGTGTTTAAACTGAGTGGAAACAGCATGAATCTCCGCTGGAAACACCTCCGAC 1137
DB 1369 ACTAATGTGTGTTTAAACTGAGTGGAAACAGCATGAATCTCCGCTGGAAACACCTCCGAC 1428
QY 1138 TATGCGGAGTAAAGGATCTGCGAATACCGCGCATCGCATCTGGAAGCCGAGCTGCTG 1197
DB 1429 TATGCGGAGTAAAGGATCTGCGAATACCGCGCATCGCATCTGGAAGCCGAGCTGCTG 1488
QY 1198 ATGTACACAGTGGGATGAGGATTTGA CGGCACTTACAGACGAAAGCTGGTGGTCGG 1257
DB 1489 ATGTACACAGTGGGATGAGGATTTGA CGGCACTTACAGACGAAAGCTGGTGGTCGG 1548

QY 1258 AACACGGCTCGTGTCTATATAGTTCCGCGCGGGATCTTCAAGTCGACGTGCAAGATCGAC 1317
DB 1549 AACACGGCTCGTGTCTATAGTTCCGCGCGGGATCTTCAAGTCGACGTGCAAGATCGAC 1608
QY 1318 ATCAGTGTGTTCCCTTCCGATGACAGCGGTGCGAGATGAAGTTGCGGAGTTGGAACCTAC 1377
DB 1609 ATCAGTGTGTTCCCTTCCGATGACAGCGGTGCGAGATGAAGTTGCGGAGTTGGAACCTAC 1668
QY 1378 GACGATTCAGCTGGATTTTACAATTAACAAGTGAACCTGGCGGTGATATCAGCAGTTAC 1437
DB 1669 GACGATTCCTCA --- 1679
QY 1438 GTGCTCAACGGGAGTGGGAACCTACTGGGTGTCGCGGCAACGTAACGAGATCTATTAC 1497
DB 1680 ---GGTGTGCCCGGCAACGTAACGAGATCTATTAC 1712
QY 1498 AACTGCTGCCCGGAACCCCTATATAGCATACCTTTCGCGCATCATCATCCGCGGCAACA 1557
DB 1713 AACTGCTGCCCGGAACCCCTATATAGCATACCTTTCGCGCATCATCATCCGCGGCAACA 1772
QY 1558 CTGTACTATTTCTTCAACCTGATCATACCTTGTGTACTGATTGCTTCCATGGCTTGTCTC 1617
DB 1773 CTGTACTATTTCTTCAACCTGATCATACCTTGTGTACTGATTGCTTCCATGGCTTGTCTC 1832
QY 1618 GGATTCACCTGCGCGCAGATTGGGTGAAAAATTTATCGCTGGGTGTTTACCATCTTGTCTC 1677
DB 1833 GGATTCACCTGCGCGCAGATTGGGTGAAAAATTTATCGCTGGGTGTTTACCATCTTGTCTC 1892
QY 1678 TCGCTGACCGGTGTTTCTGAATATGTTCCGAGACAAATGCGGCTACTTCCGATGCGGTG 1737
DB 1893 TCGCTGACCGGTGTTTCTGAATATGTTTCCGAGACAAATGCGGCTACTTCCGATGCGGTG 1952
QY 1738 CCATTG --- 1743
DB 1953 CCAITGCTGGGTACATATTTCAATTGCAATATGTTTATGTTAGTTCATCCGTTGTGTCA 2012
QY 1744 ---TG 1746
DB 2013 ACGATTTTAAATTAATTAATCATCATCGAAATGCTGATACGACGAAATGTCGGAATGG 2072
QY 1747 ATACGATCGTGTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1806
DB 2073 ATACGATCGTGTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2132
QY 1807 CCGTGATCTCTAGAGTTCCGACACACGCTGTTTCGGACACATCTCTCGAGCGGAAAGCAC 1866
DB 2133 CCGTGATCTCTAGAGTTCCGACACACGCTGTTTCGGACACATCTCTCGAGCGGAAAGCAC 2192
QY 1867 CAGATCTCTCGAGCTGAGCTGAAAGAGCGCTGCTGGAATCGTGTGCTGCTGCTGCTGCTGCTG 1926
DB 2193 CAGATCTCTCGAGCTGAGCTGAAAGAGCGCTGCTGGAATCGTGTGCTGCTGCTGCTGCTGCTG 2252
QY 1927 CTAGACATCGATGATGATTTCCGCGCAAAATTTGTCGCCCCATGACGCGCGGGAACACTG 1986
DB 2253 CTAGACATCGATGATGATTTCCGCGCAAAATTTGTCGCCCCATGACGCGCGGGAACACTG 2312
QY 1987 CCACAAACCGGCTTCTATGCGACCGTTTATGGACAAGCGCAGTGGCAGCAATTTGG 2046
DB 2313 CCACAAACCGGCTTCTATGCGACCGTTTATGGACAAGCGCAGTGGCAGCAATTTGG 2372
QY 2047 CCAATTGGCAGCAGCAGGATGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 2106
DB 2373 CCAATTGGCAGCAGCAGGATGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 2432
QY 2107 ACTGAATATGAATGAGTTTAAATCTTAAAGGAAATTCGCTTTTAACTGATCGACTACGT 2166
DB 2433 ACTGAATATGAATGAGTTTAAATCTTAAAGGAAATTCGCTTTTAACTGATCGACTACGT 2492
QY 2167 AAAGATGACGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2226
DB 2493 AAAGATGACGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2552
QY 2227 AGACTGTCCTTATCATATTTCAATGTTTCGCAATATTATGCCCAATATGCTACTACTA 2286

Qy 464 ----- 463

Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusser, D.R., Pacle, J.M., Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Stiden-Kiamos, I., Stapleton, M., Strong, R., Sun, E., Swirski, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A., Weinstein, G.M., Weissbach, J., Williams, S.M., Woodgett, W., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
The genome sequence of *Drosophila melanogaster*
Science 287 (5461), 2185-2195 (2000)
20196006
10731132

2 (bases 1 to 272521)
Celniker, S.E., Adams, M.D., Kronmiller, B., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., Benson, J., An, H., Baldwin, D., Banon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champe, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferreira, S., Frise, E., Gallie, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacle, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Richards, R., Schaefer, F., Shouanavong, S., Pittman, G.S., Puri, V., Richards, B., Scheeler, F., Stapleton, M., Strong, R., Swirski, R., Tector, C., Tyler, D., Williams, S.M., Zaveri, J.S., Smith, H.O., Venter, J.C. and Rubin, G.M.
Sequencing of *Drosophila melanogaster* genome
Unpublished

3 (bases 1 to 272521)
Misra, S., Crosby, M.A., Matthews, B.B., Bayraktaroglu, L., Campbell, K., Hradscky, P., Huang, Y., Kaminker, J.S., Prochownik, S.E., Smith, C.D., Tupy, J.B., Bergman, C.M., Berman, B.P., Carlson, J.W., Celniker, S.E., Clamp, M.E., Drysdale, R.A., Emmert, D., Frise, E., de Grey, A.D.N.J., Harris, N.L., Kronmiller, B., Marshall, B., Millburn, G.H., Richter, J., Russo, S., Searle, S.M.J., Smith, E., Shu, S., Smutniak, F., Whitfield, E.J., Ashburner, M., Gelbart, W.M., Rubin, G.M., Mungall, C.J. and Lewis, S.E.
Annotation of *Drosophila melanogaster* genome
Unpublished

4 (bases 1 to 272521)
Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA
6 (bases 1 to 272521)
FlyBase
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA
6 (bases 1 to 272521)
FlyBase
Direct Submission
Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA
On Sep 16, 2002 this sequence version replaced gi:7298121.
Location/Qualifiers
1. 272521
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="2L"
/note="genotype: Y[1]; cn[1] bw[1] sp[1]; Rh6[1]"
2745. 4311
/locus tag="CG31765"
/notes="synonym: CG8855; last curated on Thu Nov 07 12:45:45 PST 2002"
/map="34ES-34E5"

gene

/db_xref="FLYBASE:FBgn0051765"
join(2745. 2799,2858. 2975,3043. 3746,3914. 4311)
/locus tag="CG31765"
/product="CG31765-RA"
/db_xref="FLYBASE:FBgn0051765"
join(2745. 2799,2858. 2975,3043. 3746,3914. 4311)
/locus tag="CG31765"
/note="CG31765 gene product"
/codon_start=1
/product="CG31765-PA"
/protein_id="AAN10860.1"
/db_xref="GI:22946455"
/db_xref="FLYBASE:FBgn0051765"
/translation="MPDICSCEGYIGKHTQRCDDHWRGLDCKNLCCQGAACDNKSLCHCIAGWTGQFCFLPQGTGIMCKACDCDEKPCNPQTGACIQDQPLQNLVSHIVETVNSTLTKMGIIIPRTTPVLPPEVIKOPTSNENAHGPKIIVHSSSLLLEMLLLVAIAVSLVYVRYVHHKNAVYNGTITLTPANPEVLTEAAVGLVTLVSLPEPFAVTPQSENAQPELYDSFNNSSIKTPYAYARKESLYSVVSKRSGSLDRLHVTDLRIHQHQQHPRSHQHVSHSTASQFLPARPQIMPANDQFSNGSVHPHQRSHTVHLIIPPPNSNLFQVPAQITAKRAHL"
5340. 6882
/gene="BG:DS00180.12"
/locus tag="CG33119"
/note="synonym: CG8855; last curated on Thu Nov 07 12:47:27 PST 2002"
/map="34ES-34E5"
/db_xref="FLYBASE:FBgn0027929"
join(5340. 5532,5693. 6250,6329. 6532,6591. 6882)
/gene="BG:DS00180.12"
/locus tag="CG33119"
/product="CG33119-RA"
/db_xref="FLYBASE:FBgn0027929"
join(5340. 5532,5693. 6250,6329. 6532,6591. 6882)
/gene="BG:DS00180.12"
/locus tag="CG33119"
/note="BG:DS00180.12 gene product"
/codon_start=1
/product="CG33119-PA"
/protein_id="AAF53360.2"
/db_xref="GI:22946456"
/db_xref="FLYBASE:FBgn0027929"
/translation="MHCSGFIAGLTALLTVAPVQIKTSTVSTELYGDIENTEYGD FESLQSRKQHLCHREVPSPVFGNGSTIYFHRIEVCAGYRDPYAN ECPVDCSASPNDRCNGFCSPGVCEFAEFVRNEHGAHTCEIAQHGRCYLNGTC VHQNFVLDQTRQFCFKGCSQSCHECVAPQCDPCSPYRRTPLDQCPVCPADC GFGKVAPOQCECFAGIKRPNVNCVCAEVLNCEGLCSRYKCHREGYRYDVTIT SCLEPCSDNGCGNGVCIAPGVCRCFEGYVHGAECFKCERFCGKRGRCVAPICGC GEGQHCNGSCDDIEHCSPSGETHFIDRLKADRLSQHLNISEKRRHFRQLAYEF NALIGRLFN"
complement(6800. 8816)
/gene="BG:DS00180.7"
/locus tag="CG31839"
/notes="synonym: CG16882; last curated on Mon Nov 04 13:31:36 PST 2002"
/map="34ES-34E5"
/db_xref="FLYBASE:FBgn0028543"
complement(join(6800. 7190,7269. 7805,8300. 8816))
/gene="BG:DS00180.7"
/locus tag="CG31839"
/product="CG31839-RA"
/db_xref="FLYBASE:FBgn0028543"
complement(join(6881. 7190,7269. 7805,8300. 8718))
/gene="BG:DS00180.7"
/locus tag="CG31839"
/note="BG:DS00180.7 gene product"
/codon_start=1
/product="CG31839-PA"
/protein_id="AAN10861.1"
/db_xref="GI:22946457"
/db_xref="FLYBASE:FBgn0028543"
/translation="MBSSSCQVLTGLVILAI CSLGGQFKTAGIKTRPPSPGNLQIAG NSSESQWRSYNQSYYGWSYQNSQNSYANQNNQNVHVEQSGAGFVRAEVFPVTLPLPYGHY

VOPVTPPAHRVQVLDLTFKINTRSAWAGVCYKEVPTASLLRNSDQFVNGTTPD MSRLOVCDGVERPHIYRCEPICADCRNGICTANTVCVCPHVRTAEGKCIStC PLGNGVCDERNECKREGSLFETPKYQCPCKGCGSFRCAVKNKACLGDYRL AADGSPFVDSCEKDPKDLPLGNGVGVGNQCDKGTGYRDEHORNTCPHCPQGCQ ASGFWRKSAECLPKDPLGNGVGVGNQCDKGTGYRDEHORNTCPHCPQGCQ GYCAPNFCICRPGFIKSGIKRGOTCQAV" complement (10072. .11767) /gene="BG:DS00180.8" /locus_tag="CG33115" /note="synonym: CG16882; last curated on Mon Nov 04 13:14:17 PST 2002" /map="34E5-34E5" /db_xref="FLYBASE:FBgn0028542" complement (join(10072. .10629,10691. .11254,11413. .11767)) /gene="BG:DS00180.8" /locus_tag="CG33115" /product="CG33115-RA"		Query Match Best Local Similarity 33.6%; Score 776.2; DB 3; Length 272521; Matches 902; Conservative 0; Mismatches 13; Indels 98; Gaps 2;	
Qy	1	ATGAAATGACAACTGAACTGACTGAAGTTGACGATGATGAGCTGGCTGCAGTA	60
Db	103405	ATGAAATGACAACTGAACTGACTGAAGTTGACGATGATGAGCTGGCTGCAGTA	103464
Qy	61	AGATTAGCGACTGACGAGCACTTTAGCAGCTAGCAGCAAGAACACACGAGCAGC	120
Db	103465	AGATTAGCGACTGACGAGCACTATGACGAGCTAGCAGCAAGAACACACGAGCAGC	103524
Qy	121	RACAGAGGCAACACGAGCACTCAACACTGCAACCAAGGAGCTTAAAGTACAAAAAC	180
Db	103525	RACAGAGGCAACACGAGCACTCAACACTGCAACCAAGGAGCTTAAAGTACAAAAAC	103584
Qy	181	CACAGCAATTCGAAGGAGCAGCAATATGCAAGCAACAGAGCGAGCATCGAAGAC	240
Db	103585	CACAGCAATTCGAAGGAGCAGCAATATGCAAGCAACAGAGCGAGCATCGAAGAC	103644
Qy	241	GAGGATGAGCAACACGAGTGAAGCAATGACGAGCAGCAGCATCTGCAACAGCTAGAC	300
Db	103645	GAGGATGAGCAACACGAGTGAAGCAATGACGAGCAGCAGCATCTGCAACAGCTAGAC	103704
Qy	301	AGCAGCAATCTTGTGCGCAAGAGCAGCGCGAGCAACTGCTGCCGCGCATGAAGCA	360
Db	103705	AGCAGCAATCTTGTGCGCAAGAGCAGCGCGAGCAACTGCTGCCGCGCATGAAGCA	103764
Qy	361	ACAACCCAAACACCAACAAATAGACTGTGTGACGCAAGCAGCAACGATTCGTCGC	420
Db	103765	ACAACCCAAACACCAACAAATAGACTGTGTGACGCAAGCAGCAACGATTCGTCGC	103824
Qy	421	CGACGAAAGAAACACGACCAACCCCAACGAAACAGATATCA	463
Db	103825	CGACGAAAGAAACACGACCAACCCCAACGAAACAGATATCA	103884
Qy	464	-----	463
Db	103885	ATCATATTTAACTAATATATAGAAATTTAGAAATAATTTGCACCTCAGCAGCTGAAT	103944
Qy	464	-----AGAAACACAGCACTTAGCATGCTCCCTTCAAAACCGCA	505
Db	103945	TGCTCTTCTTAAATTTGCAAGAAACACAGCAACTTAGCATGCTCCCTTCAAAACCGCA	104004
Qy	506	AATCCAGGACACCTACACACACGAGCAGCAACACGAGCTGTCGACAGCAACCTACA	565
Db	104005	AATCCAGGACACCTACACACACGAGCAGCAATACAGCTGTCGACAGCAACCTACA	104064
Qy	566	TGCAATGTCGAGCAGCAGCAATGAGTTTCAATATTCGATATCGAGCATGATAGAGTAT	625
Db	104065	TGCAATGTCGAGCAGCAGCAATGAGTTTCAATATTCGATATCGAGCATGATAGAGTAT	104124
Qy	626	CCAGCGCCATTCGCTGGTGTTCATGCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	685
Db	104125	CCAGCGCCATTCGCTGGTGTTCATGCTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	104184

Qy	686	AGTGGCAACTTCACGTGCAACAGCGATCGGTGCTACTGTTTTCAGAGGATTCGAGCGACA	745
Db	104185	AGTGGCAACTTCACGTGCAACAGCGATCGGTGCTACTGTTTTCAGAGGATTCGAGCGACA	104244
Qy	746	CCATCGGCTTCATTTCCCTATTATTTAGGAGCTTTGAGCGCAACTGAAATATGAGCAGC	802
Db	104245	CCATCGGCTTCATTTCCCTATTATTTAGGAGCTTTGAGCGCAACTGAGGAGCAGCAGTA	104304
Qy	803	GCAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTA	862
Db	104305	GCAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTA	104364
Qy	863	ATAAACAATCATGATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	915
Db	104365	ATAAACAATCATGATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	104417
RESULT 9			
DROSADH02			
LOCUS			
DEFINITION			
ACCESSION			
DROSADH02			
Drosophila melanogaster, chromosome 2L, region 34C4-36A7 (Adh)			
region1, section 2 of 10 of the complete sequence			
AE003408 AC001659 L39628 L39633 L43448 L42039 L39621 L42040 L43450			
L42042 L39627 AC001302 AC000577 L42043 L39622 L39630 AC000578			
L39635 L42041 L39637 L39636 L39624 L39631 L39625 L39634 L39623			
L43449 L39626 L39632 AC001660 L39751 L49183 L42072 L39753			
L39757 L39758 L42069 L39763 L42071 L39752 L39750 L39756 L42076			
L39764 L39761 L49185 L42075 L39765 L43452 L42073 L39762 L39760			
L43451 L39754 L42068 L49184 L39755 L42074 L42070 L39759 AC001662			
L43488 L46890 L46894 L46893 L46892 L46891 L43487 AC001664 AC001202			
AC000608 AC001303 AC000611 AC001304 AC000609 AC000607 AC000613			
L43473 L79866 L43474 L42044 AC000592 L76999 AC004326 L43425			
AC000556 AC001301 AC000557 AC000558 AC000559 AC000560 AC000561			
AC000562 AC000563 AC002865 AC000564 AC000565 AC002866 AC002867			
L43424 AC002868 L39682 AC002869 L39684 L43423 AC000566 L39683			
AC000567 AC003871 AC004360 AC000622 L77011 AC000623 AC000624 L81438			
L77010 AC000625 AC001796 L81440 L81336 L81334 AC001797 AC000626			
L77009 L81439 L81335 AE003406			
HTG.			
2 of 10			
Drosophila melanogaster (fruit fly)			
Drosophila melanogaster			
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
Ephydroidea; Drosophilidae; Drosophila.			
1 (bases 1 to 320754)			
Ashburner,M., Miska,S., Roote,J., Lewis,S.E., Blazer,R., Davis,T.,			
Doyle,C., Galle,R., George,R., Harris,N., Hartzell,G., Harvey,D.,			
Hong,L., Houston,K., Hoskins,R., Johnson,G., Martin,C.,			
Moshrefi,A., Palazzolo,M., Reese,M.G., Spradling,A., Tsang,G.,			
Wan,K., Whitelaw,K., Celniker,S. and Rubin,G.M.			
An exploration of the sequence of a 2.9-Mb region of the genome of			
Drosophila melanogaster: the Adh region			
Genetics 153 (1), 179-219 (1999)			
JOURNAL			
MEDLINE			
PUBMED			
99403001			
10471707			
2 (bases 1 to 320754)			
Celniker,S.E., Agbayani,A., Arcana,T.T., Baxter,E., Blazer,R.G.,			
Butenhoff,C., Champe,D.E., Galle,R., George,R.A., Harris,N.L.,			
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,			
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,			
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,			
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,			
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., and			
Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and			
Rubin,G.M.			
Direct Submission			
Submitted (08-MAR-2000) Berkeley Drosophila Genome Project,			
Lawrence Berkeley National Laboratory, MS 64-121, Berkeley, CA			
94720, USA			

COMMENT

On or before Mar 22, 2000 this sequence version replaced
 gi:2337897, gi:1945578, gi:2337896, gi:2347073, gi:2337894,
 gi:3097819, gi:895619, gi:3097823.
 Submitted by the Berkeley Drosophila Genome Project. For more
 information, visit the BDGP Web site: <http://www.fruitfly.org/> This
 is the finished sequence of 3A4C-36A7.
 The orientation of this sequence along the chromosome is left to
 right. This sequence was annotated by Sima Misra
 (simam@fruitfly.berkeley.edu) on behalf of the Berkeley Drosophila
 Genome Center. Coding sequences are predicted based on
 computational analysis, using both gene and CDS prediction programs
 and matches to other sequences. These predictions and matches have
 been evaluated by the annotators and may have been refined by hand.
 The annotators have also used their judgement about which matches
 to include in this record. The annotations about which matches
 examined in more detail from
<http://www.fruitfly.org/publications/Adh.html>
 The annotation syntax used in this record is documented at
ftp://ftp.ebi.ac.uk/pub/databases/edgp/sequence_annotation.README.v
 1.2.

FEATURES
source

```

Location/Qualifiers
1. 320754
   /organism="Drosophila melanogaster"
   /mol_type="genomic DNA"
   /strain="y; cn bw sp"
   /db_xref="taxon:7227"
   /chromosome="2L"
   /map="3A4C-36A7"
   /clone="P1 DS01368 (D100), P1 DS08249 (D25), P1 DS08284
   (D247), P1 DS00941 (D28), P1 DS08220 (D39), P1 DS00180
   (D23), P1 DS01514 (D33), P1 DS00131 (D35), P1 DS05899
   (D22), P1 DS01759 (D41), P1 DS01523 (D34), P1 DS01652
   (D43), P1 DS03792 (D31), P1 DS01068 (D37), P1 DS06238
   (D26), P1 DS08340 (D38), P1 DS04641 (D30), P1 DS01160
   (D42), P1 DS01486 (D32), P1 DS09219 (D101), P1 DS07721
   (D20), P1 DS00810 (D21), P1 DS06874 (D19), P1 DS03431
   (D102), P1 DS03144 (D106), P1 DS03323 (D127), P1 DS01219
   (D129), P1 DS00929 (D48), P1 DS04929 (D104), P1 DS03192
   (D103), P1 DS09194 (D267), P1 DS07295 (D47), P1 DS05639
   (D11), P1 DS07851 (D49), P1 DS03023 (D8), P1 DS01845
   (D7), P1 DS04862 (D107), P1 DS07108 (D51), P1 DS09217
   (D109), P1 DS02252 (D97), P1 DS00365 (D108), P1 DS07486
   (D46), P1 DS00913 (D24), P1 DS04095 (D112), P1 DS02795
   (D50), P1 DS07473 (D113), P1 DS02740 (D27), P1 DS09218
   (D114), P1 DS02780 (D23), BAC BACR48E02 (D482), BAC
   BACR44L22 (D483)"
   /clone_lib="P1 library, partial Sau3A in pMS582tet14Ad10
   and RPl-38 (Roswell Park Cancer Institute Drosophila
   melanogaster BAC library, partial EcoRI in pBACe3.6)"
   /complement(135..1376)
   /gene="adat"
   /complement(join(135..447,505..1376))
   /gene="adat"
   /note="symbol=adat; synonym=BG:DS00941.2;
   prediction=method:'genscan', version:'1.0'; Date run:
   19-Nov-98; Time: 15:48:39', score:'54.110';
   match=method:'BLASTX', version:'2.0a19MP-WashU
   [05-Feb-1998] [Build sol2.5-ultra 01:47:35 05-Feb-1998]',
   score:'329.0', desc:'GenBank: X99383:H.sapiens mRNA for
   RNA-specific editase. CDS:148..2253; PID:e254628;
   PID:gi707504.', species:'Homo sapiens'
   /codon_start=1
   /product="hypothetical protein"
   /protein_id="AAF44818.1"
   /db_xref="GI:7287780"
   /db_xref="FLYBASE:FBan99990002"
   /translation="MCDNKKPTVEIAELCLKKFESLPKTKPTANOWTILAGVFN
   RNTEACQLVSLGGTCKIGESKLCPLNGLINDSHAVLARGLRFLYQELKQDRIFH
   RNTSTLYDMQDHEVEHFLSTQPCGDACILEEQFAARAKRQGEDESMVYTGAKL
   ISLSDPMLQTCALRTKGRGSESLSCSDCIARNVNVIGQEGALLDVLISKPTIF
   SNLNFCDLQLESLEAIFKDFCRTHFRFPQRPQINIDPGRFFFSQSDWQP
   SPNGLIWSQVPEELRPEYISVNGRQGVTKKMKTSQAMLAISKYKLFUTFLBLVFN

```

gene

CDS

```

gene
CDS
mRNA
gene
CDS

```

```

PXLSEMFQQLSDPERIAYASCKLDARYQFAWRBIKEXYFLQWTKPKHELLDFNPM
NK"
1571..2560
   /gene="BG:DS00941.3"
   /join(1571..1691,1752..1912,1978..2283,2405..2560)
   /note="symbol=BG:DS00941.3; prediction=method:'genscan',
   version:'1.0'; Date run: 19-Nov-98; Time: 15:48:39',
   score:'99.140'; match=method:'BLASTX',
   version:'2.0a19MP-WashU [05-Feb-1998] [Build sol2.5-ultra
   01:47:30 05-Feb-1998]', score:'121.0',
   desc:'tremBL: g272565:MATCH TO EST AA085966 (FRAGMENT).
   organism:HOMO SAPIENS (HUMAN). dbxref:GenBank; AC004000;
   g272565; '-', species:'HOMO SAPIENS'
   /codon_start=1
   /product="hypothetical protein"
   /protein_id="AAF44819.1"
   /db_xref="GI:7287781"
   /db_xref="FLYBASE:FBan99990001"
   /db_xref="FLYBASE:FBgn0028919"
   /translation="MPKVVSRSIVCSDTKQOEYNEEKPLNIYVCLNMKALILDCTL
   EQLPLREVDNARVINANDHANKLTHNPTPRMVYIKRSRNGIKOYVKRSCSLPL
   YRHSPDSHVTVMNSALIRNKGESPLTOLLNSIKGSFKAPAKPATISAGPDSGIV
   DAGSKVVVTEHTKMGKFSVTVTIDEEDEIBAREIADSIANNARIIEKQLQKRG
   KGLSDVGIKTIKEDAPPFPKQKQRTLLER"
   /join(3739..5231,5285..5642,5705..6627,6692..7540,
   7603..8493,8562..9195,9261..9751)
   /note="symbol=Sos"
   /gene="Sos"
   /match=method:'sim4', synonym=BG:DS00941.4;
   desc:'GenBank: M83931.Drosophila melanogaster son of
   sevenless (Sos) mRNA, complete cds. CDS:346..5133;
   PID:g158485.', species:'Drosophila melanogaster';
   match=method:'BLASTX', version:'2.0a19MP-WashU
   [05-Feb-1998] [Build sol2.5-ultra 01:47:30 05-Feb-1998]',
   score:'889.0', desc:'tremBL: 007889:SON OF SEVENLESS 1
   PROTEIN (SOS 1)', dbxref:GenBank; L13857; 9306778; '-',
   species:'HOMO SAPIENS'
   /codon_start=1
   /evidence=experimental
   /protein_id="AAF44820.1"
   /db_xref="GI:7287782"
   /db_xref="FLYBASE:FBan99990000"
   /db_xref="FLYBASE:FBgn0001965"
   /translation="MFGSPGSHAHTISYGGGIGLGTGGGGSGSGSGSGGGGIGI
   GGGVAGLQDCDGYDTKCNAAWRGLFTPSLKVLEQVHPRVATAKEDALLYVEKLC
   LRLAMLCAPKPHSVQDVEEKVNSFPAPIDQWALNKAKEVINSKKRKSVLPTKRVH
   TLQKDVLOYKIDSSVSAFLVALEVISADILKMGADYVIKIAHCETIKEDIEVMNA
   DRLVMDLNOSEAHILPSLPAQASATVEETVKELHDEKQYQDRLHMIIRVFE
   EEDVYKAYLEPIFSNIMDIYEVTVILGSLIEDVIEMSGQSPACVSCFELABA
   HARVFDYIKHLDSSQDIESFEQVQLLPHLCLEKVMASIKRQVPSGRV
   RQGLAIFRELQKQVHEDKQGNCFREDISLKSGLSKGRKTSBERKFLFDGL
   MVLCKANTKQTSAGATAYRLKEKPYMRVYDINDRPSDDLKNSFELAPMQPI
   VLTAKNAOHKHDMAVLLAVITKSLDRLHLSILOIERKHLRMPSPETIKFAVDS
   GDNVILEESAGVPMIKATLCKLRLHTYIADPTFVRTITTYRIFCSPPQQLLQ
   LLVERFNIPLSVQDITAGAGMGVGGKHEKNSHREDMRKRYKQVQVPRV
   LNVLRHWDHFFYDFEKDPMLEKLFNLHLEVNGSKMRKWDSVLKIVQRKQEKSN
   KKITVAYGHDPPIIEHLSVNPDEITLLTHLELARQLTLLEFVEKNVKNPSELVGS
   PNFKDEKVSNNLIMKHTTNVTVEIKSIEAENYERLAIMQRALEVMMVLEL
   NNFNGILSIVAANGTASVYRLWTFGLPERYRKFLERELSDHDLKKIKERLIN
   PCVPFFGRLTNILHLEGNPDLLANTELFNSRKRKVAEIIGETIQYQNOYCLNE
   ESTIQFQPLDFNGLSDKSDYLYNESLIEPRCKTVPKPRKPHIPLKSPGI
   KPRQNTNSSSKLSNSTSVAASAAASATSIATASAPLSHSAIMDAPIAAANA
   GSGTLAGESPQHNSAFVAPVPIPERNTSWSCTPOHTRTDQNGSVSPAPHL
   KPGAHVANNNSTLASAMDVPSPALPEHLPSQLSDSNPPASDTEAPSPFLK
   VVSPRHETGRSPFHGRMQNSTHSTASTVTLTGSTSGEGEFCAGGFNFGAQCGQP
   GAVPISPHVNVPMATMETRAYVPPPLPPRRKERTESCADMAQKQAPDAPTLPPRDE

```


985 CCGCTTCTCAATGAATCGACCCGCTTAACTTAAGCTTTGGTTTAACTTTAATCAAAAT 1044
1045 ATCGATGTGACGAGAAATCAATGCTAGTCACTAATGTTGGTTTAACTTAACTGAGTGG 1104
1105 AACGACATGATCTCCGCTGGAACACCTCCGACTATGCGGAGTTAAGGATCTCGAATA 1164
1165 CCGCGCATCGCATCTGGAAGCCGAGCTGTGATGATGATGATGATGATGATGATGATGAT 1224
1225 GACGACCTACAGACGAAAGCTGTGTTGGGAAACAGCGCTCGTGTCTATCTGTTCCG 1284
1285 CCGGGATCTCAAGTCCGACGTCGAAGATCGACATCACTGCTGCTTCCCTTCGATGACAG 1344
1345 CCGGATCTCAAGAGACCTGCAAGATCGACATCACTGCTGCTTCCCTTCGACACCA 844
1404 CCGTCCGAGATGAAGTTCGGAGTGGACCTACGACGATTCAGCTGCTGATGATGATGATGAT 1404
1405 CAAGATGAAGTGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1464
1465 GGTGTCGCGGCAACAGTAAACAGATCTATTAACATCTGCTGCTGCTGCTGCTGCTGCTGCT 1524
1525 ATCACTTCGCGCATCATCGCGGACGAAACACTGATCTATTTCTCAACCTGATCATA 1584
1585 CTTGTGTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1644
1645 GAAATATATCGTGGTGTCTACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704
1705 GCGGACAAATCCCGCTACTTCCGATGCGGTGCCATG-----TGGATACG 1752
1753 ATCGTGTGTTTG--TGCTGGCTGCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1810
1811 TGATCTGATGTTCCGACACCGCTGTTGCGACACATCTCCCGAGCGAA-----GCAC 1866
1867 CAGATGATCTCCGACGTTGAGGTGAAGAGCGCTGCTGAAATCGCTGCTGCTGCTGCTGCTGCTGCT 1926
1927 CTAGACATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1986
1987 CCACAAACCCCGCTTCTATCGACCGTTTATGACAAAGCGGACGATGATGATGATGATGATGATGAT 2046
2047 CCAATTGGACAGACCCGAAATGCGGATGCGGTCACCCATCATACGCTGATCAAAATCATCA 2106

1565 TACAGGGGGGTGAGGAGATGGCGCGGGGTGGCGCGCACAGTTGCTTC-----GGT 1618
1619 GTCGACTACGAGCTCTCCCTCATCTCTGAGAGAGATTAGATCATCACAGATCAGATCGC 1678
1679 AAGGACGACGAAGATGCGGACATTTGCGCGACTGGAAGTTGCGCGCCATGCTCGTGAC 1738
1739 AAGTGTGCTTATATCTTACCTGTTCATCATTCGCGACCTAGCCGCTGCTGCTG 1798
1799 TCCGGCCACATCATGCTGCTG 1822
2287 TCGGCACCATATTTATGCTCTG 2310
2310 TCCGGCCACATCATGCTGCTG 1822

RESULT 12
AX009612
LOCUS AX009612
DEFINITION Sequence 3 from Patent EP0962528.
ACCESSION AX009612
VERSION AX009612.1 GI:9996844
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuidae; Noctuidae; Heliothinae; Heliothis.

1
Adamczewski, M.D., Schultze, T.D. and Oellers, N.D.
Nucleic acids encoding acetylcholin-receptor subunits from insects
Patent: EP 0962528-A 3 08-DEC-1999;
BAYER AG (DE)

FEATURES
Location/Qualifiers
1..3700
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"
335..1825
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07500.1"
/db_xref="GI:9996845"
/translation="MGRRARRSLAAPAGLLLLCLLMPRGARCGVHEKRLHLLLDH
YNLBRPVNEDPLQLSGLTLMQIDVDKRNQLITNINIKLWMDNLRWNTSDF
GSKDLRVPVHRLKRPDLVMSADEGDFSTPTNVVNRNGLSVPPGIPKSTCKI
DITWPFDDQRCMKFGSWTDYQDLQLODEGGDISSFTNGEWELIGVGRNE
IYNCCPEPYDITFAVIRKRLTYFFNLI VPCVLIASMALLGTLPPDSGKLSIG
TLLLSLTVFLNVAETMPATSDVAPLLGTPTNCIMFMVASSWSTLILNHYHRRAD
THEMSDWIRCVFLYMLPWRMSRPGSATPPPARVPPDLELRSSKSLANVLD
IDDFRHFQAGQCCRYRGENGAGLAHSCFQVDYELSLILKEIRVITDQMRKD
DEDAISRDNWKFAMVVDRLCLILFTLITATLALLSAPHMVS"

CDS

Query Match 22.2%; Score 512.8; DB 6; Length 3700;
Best Local Similarity 62.5%; Pred. No. 4.3e-129;
Matches 878; Conservative 0; Mismatches 502; Indels 24; Gaps 4;
925 GGATATCATGAAAGAGACTGTTACAGCATCTTTGGATCTTTATTAATACACTAGACGT 984
425 GGGTACACGAGAGCGGCTACTGACACCTATTGGACCACTACAACTACTGGAGAG 484
985 CCGGTTCTCAATGAATCGACCCGTTACAAATTAAGCTTTGGTTTAACTTTAATCAAAAT 1044
485 CCGGTCGTCACGAGAGCGACCGCTGCTCTCCGCTCAGCTCATGACATC 544
1045 ATCGATGTGACGAGAAATCAATGCTAGTCACTAATGTTGGTTTAACTTTAATCAAAAT 1104

QY	2287	TGGGACACATATTATTGTCTCG	2310
Db	1799	TCGCGCCACCATCATGTGTCTCG	1822
RESULT 14			
AF321445	2023 bp	mRNA	linear
LOCUS	INV 29-APR-2002		
DEFINITION	Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6 subunit variant type I (nAcRalpha-30D) mRNA, complete cds, alternatively spliced.		
ACCESSION	AF321445		
VERSION	AF321445.1	GI:20152844	
KEYWORDS	Drosophila melanogaster (fruit fly)		
SOURCE	Drosophila melanogaster		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 2023) Grauso,M., Reenan,R.A., Culetto,E. and Sattelle,D.B. Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, fya Dalphas, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing Genetics 160 (4), 1519-1533 (2002)		
JOURNAL			
MEDLINE	21969411		
PUBMED	11973307		
REFERENCE	2 (bases 1 to 2023) Grauso,M. and Sattelle,D.B. Direct Submission Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional Genetics Unit, University of Oxford, South Park Road, Oxford OX1 3QX, UK		
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source	1..2023 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /chromosome="2" /map="30DI" /dev_stage="embryo"		
gene	1..2023 /genes="nAcRalpha-30D"		
CDS	379..1863 /gene="nAcRalpha-30D" /note="ion channel; neurotransmitter transmembrane receptor; alternatively spliced; contains exons 3a and 8b" /codon_start=1 /product="nicotinic acetylcholine receptor Dalpha6 subunit variant type I" /protein_id="AAM13392.1" /db_xref="GI:20152845" /translation="MSPPLASLSFLVLLIFLAIKESCGPHEKRLNHLSTNTL ERPVANESEPLEVKFGTLTQQIIDVDENKQILTTNAWLENDYNLRNETSYGVK DLRTITNKMKPDVLMYNSADEGFDTHNIIVKNGSLYVPPIFKSTCKIDTFV FPDDQHCKMFSGWTYDGNQLDLVNSDEGGLSDFITNGEWYLAMPGRKNITIYA CSPEYVDITFTIQIRRRYLYPFNLIVPCVLTSSMALLGFTLPDPSGGEKLTVL LSVTLFNLVAETLPQVSDAIPLLGYFCNIMFMVASVVLTVVNLNYHRTADIEH PPWKTSVFQWLMPILRMGRPGRKITRKTIILSNMRKELEKRSKSLLIANVLDD DFRTISGSQTALGSSASGRPTVEEHHTAGCNHKDLHLILKELQFITARMRKAD FBSLIGDKWFAPMVDRFCLIVFTLTIIATVTVLVSAPHLIIVQ"		
variation	793..794 /genes="nAcRalpha-30D" /notes="results in asparagine to glycine substitution; compared to variant clone" /replace="gg"		
variation	846 /gene="nAcRalpha-30D" /note="results in isoleucine to methionine substitution; compared to variant clone" /replace="g"		
variation	938		

variation	/gene="nAcRalpha-30D"		variation	/note="results in asparagine to serine substitution; compared to variant clone"	
	/replace="g"			1314	
variation	/gene="nAcRalpha-30D"		variation	/note="compared to variant clone"	
	/replace="a"			1560	
variation	/gene="nAcRalpha-30D"		variation	/note="compared to variant clone"	
	/replace="t"			1605..1607	
variation	/gene="nAcRalpha-30D"		variation	/note="results in serine deletion; compared to variant clone"	
	/replace=""			1683	
variation	/gene="nAcRalpha-30D"		variation	/note="compared to variant clone"	
	/replace="t"			1683	
ORIGIN					
Query Match 19.4%; Score 448; DB 3; Length 2023;					
Best Local Similarity 58.5%; Pred. No. 2.5e-111;					
Matches 828; Conservative 0; Mismatches 570; Indels 18; Gaps 2;					
Qy	909	TAAAGTTTGGCTACGAGATATCATGAAAAGAGACTGTTACACGATCTTTGGATCCCTTA	968		
Db	441	TAAAGAAAGCTGTCAAGGACCTCATGAAGCGCCTGCTGAACCATCTGCTGTCCACCTA	500		
Qy	969	TAATACACTAGAACGTCCTTCATGATCGAGCCGTTCAAAATTAAGCTTTGGTTT	1028		
Db	501	CAATACGCTGGAGCGACCGCTGGCCAAATGAATCGAGCCCTCGAGGTTAAGTTCCGACT	560		
Qy	1029	AACCTTTAATCGAATTCGATGTGGACGAGAAAATCAATTTGTAGTCACTAATGTGTG	1088		
Db	561	GACGCTGACGAGATCATCGACGTGGATGAAAGAAATCAGATCTTGACCAAAATGCGTG	620		
Qy	1089	GTTAAAACTGGAGTGGAAACGACATGAATCTCCGCTGGAACACCTCCGACTATGCGGAGT	1148		
Db	621	GTTAAATTTGGAGTGAACGACTACATCTCGCTTGGAAATGAACGGAATACGCGGGGT	680		
Qy	1149	TAAGGATCTCGAATACCGCGCATCTGGAAGCCGAGCTGTGATGTACAACAG	1208		
Db	681	CAAGGATCTAGAAATCAGCGCCAAACAGCTGTGGAAGCCCGACGTGCTCATGTACAACAG	740		
Qy	1209	TGCGGATGAGGATTTTACGCGCACTACAGACGATGGTGGTGGGAAACACGCGCTC	1268		
Db	741	CGCGGATGAGGATTCGATGGCAGTATCACCCAAATTTGGTCAAAATACACGCGAG	800		
Qy	1269	GTGCTATAGTTCCGCGGGGATCTTCAAGTCGACGTGCAAGATCGACATCAGTGGTT	1328		
Db	801	TTGTCTGTAGTGGCCCTGGTATCTTCAAGAGCACAATGCAAGATAGACATCAGTGGTT	860		
Qy	1329	CCCTTCGATGACAGCGGTGGAGATGAAGTTGCGGAGTTGGACCTACGACGGAATCCA	1388		
Db	861	CCCATTTGATGACCAACATTTGCGAAATGAAATTCGGTAGTTGGACTTACGATGGAATCA	920		
Qy	1389	GCTGGATTTACAATTACAAGATGAACCTGGCGGTGATATCAGCAGTTACGTGCTCAACGG	1448		
Db	921	GTTGGATTTGGTTTGAATTCGGAAGATGAGGGGATCTTTCGGATTTCAACAATATGG	980		
Qy	1449	CGAGTGGAACTACTTGGGTGTGCCCGCAACGTAACGAGATCTATTACAACTGTGCGCC	1508		
Db	981	CGAGTGTACTTGTCTGCCATGCGGGAAGAGATACGATAGTCTACGCCCTGTGCGCC	1040		
Qy	1509	GGAACTTATAGATACCTTTCCGCATCATCATCCGCGACGACACTGTACTATTTT	1568		
Db	1041	AGAACCATATGTGGATATACCTTTACTATACAAATTCGTGCGCGTACATTATATTTT	1100		
Qy	1569	CTTCAACCTGATCATACCTTTGTGTAATGTCCTCCATGGCCTTGTCTCGAATTCACCT	1628		

Db	1101	TTTCAATTTAATCGTCCCATGTGTCTAATCTCATCGATGGCCCTACTGGGCTTCAATT	1160		
Qy	1629	GCOCAGAGATCGGGTGAAAAATTTATCGCTGGGTGTATACCATTTTGTCTCTCGCTACCGT	1688		
Db	1161	CCCGCGGATTCGGCGAGAAACTGACGCTGGAGTTACAATTTCTTCTATCGCTCACAGT	1220		
Qy	1689	GTTTCTGAATATGTTGCCGAGACAATGCCGGCTACTTTCCGATGCGGTGGCAATTTGGAT	1748		
Db	1221	GTTTCTCAACCTTTGTAGCTGAGACATTTGCCCAAGATATCTGATCAATTCCTCTTTAGG	1280		
Qy	1749	ACGCATCGTGTTTTGTGCTGCGCATGATATTCGAATGATGAGTCCCGAGGACGACC	1808		
Db	1281	CACCTACTTCAATTCATCATGTTCAUGSTCGCTCGTGGTGGTGTGACAGTAGTGGT	1340		
Qy	1809	GCTGATCCTAGATTTCCCGACACGCCCTGTTTCGACACATCTCTCCGAGCGGAAGCAC--	1866		
Db	1341	GCTCAACTACCCATCGCACAGCGACATTCACGAGATGCCACCGTGGATCAAGTCGT	1400		
Qy	1867	---CAGATACTCTCCGAGTTGAGCTGAAGAGCGCTGCTCGAAATCGTCTCGGCCAAC	1923		
Db	1401	TTTCTCAATGCGTCCCTGGATCTTTCGGAATGGTGTGACCCGGTTCGCAAGATTACAG	1460		
Qy	1924	GTACTAGACATCGATGATGACTTCCG-----GCACAATTTGCGCCCATGAC	1970		
Db	1461	CAAAACAATACTATTAGCAATCCATGAAGAGCTGGAGCTAAAGAGCGCTCTCTCCAA	1520		
Qy	1971	GCCCGGCGGAACACTGCCACACAAACCGCGCTTTCTATCGCACGGTTTTATGACAAAGCGA	2030		
Db	1521	ATCCCTCTGCGCAATGTCTCGACATCGACGACGCTTCGGGCACACAAATCTGGGTC	1580		
Qy	2031	CGATGGCAGCATTTGGGCAATTTGCGAGACCCGGAATGCGGATGCGGTGATCCCATATAC	2090		
Db	1581	CCAAACCGCATTTGGCTCTGTCGCGCAGCTTCGTCGCGCCCAACAACGGTGGAGGACATCA	1640		
Qy	2091	GTGCATCAATCACTCAACTGAATATGAATAGTATTAGTTTAACTTAAAGGAATTCGCTTAT	2150		
Db	1641	CACGGCATCGGTGCAATCACAAGATCTTCACTCTAATTTCTCAAGAATTTGAATTTAT	1700		
Qy	2151	AACGTGATCAGTACGTAAAGATGACGAGTGCATGCAATTTGCCAATGATTGGAAATTTGC	2210		
Db	1701	TACGCGCGGATCGCAAGCTGACGACGAAGCGGAATTTGATCGCGCATTTGGAAGTTGCG	1760		
Qy	2211	AGCTATGCTGTTGACAGACTGCTTATCATATTTACAAATGTTCCGAATTTAGCCAC	2270		
Db	1761	GGCAATGGTTGTGATAGATTTTGTAAATTTGTTTTCACGCTCTTACGATTTATGCAAC	1820		
Qy	2271	AATAGCTGTACTACTATCGGCACCAATATTATTGT 2306			
Db	1821	GGTTACGGTCTGCTCTCCGCTCGGCACATAATCGT 1856			
RESULT 15					
AF321447					
LOCUS					
DEFINITION					
Drosophila melanogaster nicotinic acetylcholine receptor Dalphae					
subunit variant type III (nAcRalpha-30D) mRNA, complete cds,					
alternatively spliced.					
ACCESSION					
AF321447					
VERSION					
AF321447.1 GI:20152848					
KEYWORDS					
SOURCE					
ORGANISM					
Drosophila melanogaster (fruit fly)					
Drosophila melanogaster					
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;					
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;					
Ephydroidea; Drosophilidae; Drosophila.					
1. (bases 1 to 2023)					
REFERENCE					
AUTHORS					
TITLE					
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,					
Dalphas, Dalphae and Dalpha7, in Drosophila melanogaster Identify a					
New and Highly Conserved Target of Adenosine Deaminase Acting on					
RNA-Mediated A-to-I Pre-mRNA Editing					
JOURNAL					
Genetics 160 (4), 1519-1533 (2002)					
MEDLINE					
21969411					

2023 bp mRNA linear INV 29-APR-2002

Drosophila melanogaster nicotinic acetylcholine receptor Dalphae subunit variant type III (nAcRalpha-30D) mRNA, complete cds, alternatively spliced.

AF321447

AF321447.1 GI:20152848

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1. (bases 1 to 2023)

Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B. Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas, Dalphae and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing

Genetics 160 (4), 1519-1533 (2002)

21969411

PUBMED REFERENCE AUTHORS TITLE JOURNAL	11973307 2 (bases 1 to 2023) Grauso,M. and Sattelle,D.B. Direct Submission Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional Genetics Unit, University of Oxford, South Park Road, Oxford OX1 3OX, UK	Location/Qualifiers 1. .2023 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /chromosome="2" /map="30D1" /dev_stage="embryo" 1. .2023 /gene="nAcRalpha-30D" 379. .1863 /gene="nAcRalpha-30D" /note="ion channel; neurotransmitter transmembrane receptor; alternatively spliced; contains exons 3b and 8b" /codon_start=1 /product="nicotinic acetylcholine receptor Dalpha6 subunit variant type III" /protein_id="AA013394.1" /db_xref="GI:20152849" /translation="MDSPLPASLSFVLLIFLAIIKSCQGPHEKRLNHLSTVNTL ERPVAISEPLEVKGFLTQQIIDLVDKKNLLITNLSLEWMDYNLRNTEREYGVK DLRITPKMLKPVLYNSDEGDFGTHTNIVKHSGLCYVPGPIFKSTCKMDITV FPFDHCECMKFSWTYDGNLDVLNSDEGDLSDFITNGEWYLLAMPKKNITVYA CCPEPYVDITQIRRTLYFFNLIVPCVLISSMALLGFTLPDPSGEKLTGLVITL LSLTVNLVAETLPOVSDAIPLLGTIVFCIMPVASSVLTVVLNVYHRTADIEH PMTKSVFLOMLPILMGRPRGKTRKTIILSNRMKELEKRSKSLANVLDDIDD DFRHTISGQTALGSSASGRPTVEEHHTAIGCNKDLHLILKELQPIFTARMKADD BAELIGDWKFAAMVVDRELFVTLFTIIATVTVLLSAPHIIIVQ" 794	variation /gene="nAcRalpha-30D" /note="results in glycine to serine substitution; compared to the sequence deposited in GenBank Accession Number AE003626" /replace="a" 846	variation /gene="nAcRalpha-30D" /note="results in asparagine to serine substitution; compared to the sequence deposited in GenBank Accession Number AE003626" /replace="a"	ORIGIN Query Match Best Local Similarity 58.2%; Score 441.6; DB 3; Length 2023; Matches 824; Conservative 0; Mismatches 574; Indels 18; Gaps 2; QY 909 TAAAGTTTCCCTAGCAGATATCATGAAAGAGACTGTGTACAGATCTTTGGATCCTTA 968 DB 441 TAAAGAAAGCTGTCAAGACCTCATGAAAGCGCTGTGTAACCATCTGCTGCCACCTA 500 QY 969 TAATACACTAGAACGTCCTCGTTCTCAATGAATCGGACCGGTTCACATTAAGCTTTGGTTT 1028 DB 501 CAATACGCTGGAGCGACCGGTGGCCCAATGAATCGGAGCCCTCGGAGGTTAAGTTCGGACT 560 QY 1029 AACTTTTAATGCAAAATTCGATGTGGACGAGAAAAATCAATGCTAGTCACTAATGTGTG 1088 DB 561 GACGCTGCAGCAGATCATCGACGTGGACGAGAAGAAATCAGCTTCTCATAAACGAATCTTTG 620 QY 1089 GTTAAACCTGGATGGAGCAGCATGAATCTCCGCTGGAAACACCTCCGACTATGGCGGAGT 1148 DB 621 GCTTTCGTTGGATGGAAAGACTACAAATCTCGCTGGAATGAACGGAATACGGCGGGT 680 QY 1149 TAAGGATCTGCGAATACCGCGCATCGCATCTGGAAGCGGACGCTGCTGATGTACAAACAG 1208 DB 681 CAAGGATCTACGAATCACGCCCAACAGCTGTGGAGCCCGACGCTCATGTATGTAACAACAG 740
QY	1209	TGCGGATGAGGATTTGACGCGCACCTACACAGCAAGCTGTGTGTGCGGACAAACAGGCTC	1268		
DB	741	CGCGGATGAGGATTCGATGGCAGCTATCACACCAACATTGTGTGTTCAACATAGCGGAG	800		
QY	1269	GTGTCTATACGTTCCGCGGGGATCTTCAAGTCAAGCTGCAAGATCGACATCAGTGTGTT	1328		
DB	801	TTGTCTGTACGTGCCCCCTGGTATCTTCAAGAGCACATGCAAGATGGAATCAGCTGTT	860		
QY	1329	CCCCTTCGATGACGAGCGGTGCGAGATGAAGTTCCGCGAGTTGGACCTACGACGGAATCCA	1388		
DB	861	CCCAITTTGATACCAACCAATTCGGAATGAATAATTCGTAGTTGGATACGATGGAAATCA	920		
QY	1389	GCTGGATTTACAATTAACAAGATGAACCTGCGGCTGATATACGAGTTACGCTCAACGG	1448		
DB	921	GTTCGATTTGGTTTGAATTCGGAAGATGGAGGGGATCTTTCGATTTTCAACAAATGG	980		
QY	1449	CGAGTGGGAACACTACTGGGTGTGCCCGGAAACGTAACGAGATCTATTAACAACTGTGCCC	1508		
DB	981	CGAGTGGTACTTGTCTTGCCATGCCGGAAGAAATACGATAGTCTACGGCTGTCTGCC	1040		
QY	1509	GGAACCTTATATAGACATCACTTCGCCCATCATCCGCGGACGAAACACTGTACTATTT	1568		
DB	1041	AGAACCATATGTCGATATCACCTTTACTATACAAATTCGTGCCGTACATTATATTTT	1100		
QY	1569	CTTCAACCTGATCATACCTTGTGTACTGAITGCCCTCCATGCGCTTGTCTCGGATCACCT	1628		
DB	1101	TTTCAATTTAATCGTGCCATGTGTCTAATCTCATCGATGGCCCTACTGGGCTTCACAT	1160		
QY	1629	GCCTGAGATTCGGGTGAAAAATATTCGTGGGTGTATACCATCTTGCTCTCGCTACCGT	1688		
DB	1161	GCCTGCGGATTCGGGCGGAAACATGACGCTGGGAGTTACAATCTTCTATCGCTCACAGT	1220		
QY	1689	GTTCCTGAATATGGTTTCCGAGACAAATGCCGCTACTTCCGATGCGGTGCCATTTGGAT	1748		
DB	1221	GTTCCTCAACCTGTAGCTGAGACATTCGCCCAAGTATCTGATGAATCCCTCTGTAGG	1280		
QY	1749	ACGATCGTGTTTTTGTGTGCTGCCATGGAATATTTGGAATAGTGTGCCCGGACGAGCAC	1808		
DB	1281	CACCTACTTCAATTTGCATCATGTTTCATGTCGCCCTCGTGGTGTGCTGACAGTAGTGT	1340		
QY	1809	GCTGATCCTAGAGTTCCTCGACCAACGCCCTGTTCGGACACATCTCCGAGCGGAAGCAC--	1866		
DB	1341	GCTCAACTACCACCATCGCAGCGGACATTCACGAGATGCCACCGTGATCAAGTCCGT	1400		
QY	1867	---CAGATACTCTCCGACGCTTGAGCTGAAAGAGCGCTCGTGAAATTCGCTGTGGCCAA	1923		
DB	1401	TTTCTTACAATGGCTGCCCTGGATCTTGGGAATGGTGCACCCGCTCGCAAGATTACACG	1460		
QY	1924	GTACTAGACATCGATGACTTCGCG-----GCACATTTGTCGCCCATGAC	1970		
DB	1461	CAAAAACAATACTATTAAAGCAATCGCATGAAGAGCTGGAGCTTAAAGAGCGCTCTCCAA	1520		
QY	1971	GCCCGCGGAACAACACTGCCACAAACCCGCTTCTATCGACGGTTTATGGACAAGGCGA	2030		
DB	1521	ATCCCTGTGGCCAAATGTCCTGACATCGACGACGACTTCGGGACACAAATCTTGGCTC	1580		
QY	2031	CGATGGCAGCATTTGGCCAAATTTGGCAGCACCCGGAATGCGGATGCGGTCAACCATCAT	2090		
DB	1581	CCAAACCCCATTTGCTCGTCCGCGCAGCTTCGTGCGGCCCAACACGTTGGAGGAGCATCA	1640		
QY	2091	GTGCATCAAAATCATCAACTGAATGAATTAGTTTAACTTTAAAGGAAATTCGCTTTAT	2150		
DB	1641	CACGGCCATCGGCTGCAATCAAAAGATCTTCATCTAATTTCTCAAGAAATTCGAATTTAT	1700		
QY	2151	AACTGATCAGCTACGTTAAAGATGACGAGTGCATGCAATGCAATTCGCAATGATGGAATTTGC	2210		
DB	1701	TACGGCGGATGCGCAAGCTGACGACGAGCGGAATGATCGGCGAATTTGGAAGTTCG	1760		
QY	2211	AGCTATGTTGCTTGACAGACTGTGCTTATCATATTTCAAAATTTTGCATATTTAGCCAC	2270		
DB	1761	GGCAATGGTGTGGATAGATTTTGTAAATTTTAAATTTTTCACGCTCTTCACGANTATTGCA	1820		
QY	2271	AATAGCTGCTACTACTATCGGACCAACATATTTGT	2306		

Db 1821 GGTACGGTGTGCTTCCTCCGCGACATATCGT 1856

Search completed: May 8, 2004, 02:41:13
Job time : 9171.23 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 10:08:00 ; Search time 929.055 Seconds
(without alignments)
10562.710 Million cell updates/sec

Title: US-09-303-232-1_COPY_372_2681

Perfect score: 2310

Sequence: 1 atgaaaaatgcacactgaa.....caccacatatattgtctcg 2310

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2310	100.0	2886	3	Aaz24475 D. melano
2	881	38.1	984	4	Abi27131 Drosophil
3	776.2	33.6	3144	4	Abi27130 Drosophil
4	531	23.0	803	4	Abi07799 Drosophil
5	512.8	22.2	3700	3	Aaz24476 H. viresc
6	479	20.7	936	4	Abi13733 Drosophil
7	412.2	17.8	1540	4	Abi07231 Drosophil
8	411.4	17.8	3109	3	Aaz24477
9	335.8	14.5	14668	4	Abi07798 Drosophil
10	296.2	12.8	1509	3	Aac58395 Human PRO
11	296.2	12.8	1509	4	Aac90380 Wild-type
12	296.2	12.8	1876	2	Aav12197 Human neu
13	296.2	12.8	1876	2	AAT48239
14	296.2	12.8	1876	6	ABSS4875
15	296.2	12.8	1876	6	ABV73248
16	296.2	12.8	1876	8	ADAI0864
17	294.6	12.8	1509	4	AAC90385
18	294.6	12.8	1509	4	AAC90386
19	294.6	12.8	1590	2	AAV44687
20	293	12.7	1509	2	AAC90387
21	284.4	12.3	2101	2	AAT59197
22	283.8	12.3	1964	6	ABZ11298
23	267.8	11.6	2769	2	AAT59196

24	251	10.9	2277	2	AAV12199	Aav12199 Human neu
25	251	10.9	2277	2	AAQ90387	Aac90387 Alpha 2 s
26	251	10.9	2277	6	ABSS4870	ABSS4870 Human neu
27	251	10.9	2277	6	ABV73243	Abv73243 Human neu
28	251	10.9	2664	2	AAT48235	Aat48235 Neuronal
29	251	10.9	2664	6	ABK92165	Abk92165 Prostate
30	251	10.9	2664	8	ADAI0854	Adai0854 Human neu
31	251	10.9	2664	9	ADC71170	Adc71170 Human 205
32	251	10.9	2666	9	ADC71168	Adc71168 Human 205
33	245.6	10.6	1416	4	AAC90382	Aac90382 Chimeric
34	242.6	10.5	1809	5	AAS91552	Aas91552 DNA encod
35	242.6	10.5	2082	2	AAT59527	Aat59527 Alpha4 su
36	242.6	10.5	2752	9	ADB78668	Adb78668 Human nic
37	242.6	10.5	2752	9	ADB78671	Adb78671 Human nic
38	242.6	10.5	2752	9	ADB78670	Adb78670 Human nic
39	242.6	10.5	2752	9	ADB78661	Adb78661 Human nic
40	242.6	10.5	2752	9	ADB78672	Adb78672 Human nic
41	242.6	10.5	2752	9	ADB78663	Adb78663 Human nic
42	242.6	10.5	2752	9	ADB78669	Adb78669 Human nic
43	242.6	10.5	2752	9	ADB78662	Adb78662 Human nic
44	242.6	10.5	2752	9	ADB78673	Adb78673 Human nic
45	242.6	10.5	3343	9	ADE85024	Ade85024 Farnesyl

ALIGNMENTS

RESULT 1
AAZ24475
ID AAZ24475 standard; CDNA to mRNA; 2886 BP.
XX
AC AAZ24475;
XX
DT 17-FEB-2000 (first entry)
XX
DE D. melanogaster acetyl-choline receptor DNA from clone Da7.
XX
KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
XX neurotransmission; plant protection agent; conductance; AChR; ds.
XX
OS Drosophila melanogaster.
XX
PH Key Location/Qualifiers
CDS 372..2684
FT /*tag= a
FT /product= "acetyl choline receptor."
XX
PN DE19819829-Al.
XX
PD 11-NOV-1999.
XX
PF 04-MAY-1998; 98DE-01019829.
XX
PR 04-MAY-1998; 98DE-01019829.
XX
PA (FARB) BAYER AG.
XX
PI Adamczewski M, Oellers N, Schulte T;
XX WPI: 2000-014207/02.
DR P-PSDB; AAY50814.
XX
PT New nucleic acid encoding a nicotinic acetylcholine receptor from
XX insects, used to identify potential insecticides.
XX
PS Claim 1a; Page 8-12; 26pp; German.
XX
CC This invention describes a novel nucleic acid (NA) encoding a nicotinic
CC acetyl-choline receptor (I) from insects which can be used as an
CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
CC (also vectors containing it, its regulatory regions, and antibodies
CC directed against (I)-encoded proteins) are used to screen for: (a) plant
CC protection agents that alter conductance of AChR, potentially useful as

DR P-PSDB; ABB63696.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT Genes from *Drosophila* and for elucidating cell signaling and cell-cell

PT interactions.

XX

XX Claim 1; SEQ ID NO 17879; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB572072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 803 BP; 211 A; 199 C; 190 G; 203 T; 0 U; 0 Other;

Query Match 23.0%; Score 531; DB 4; Length 803;

Best Local Similarity 88.4%; Pred. No. 1.8e-138;

Matches 616; Conservative 0; Mismatches 15; Indels 66; Gaps 1;

QY 1054 GACGAGAAAATCAATTGCTAGTCACTAATGTGTGGTTTAAACTGGAGTGAACGACATG 1113

Db 1 GACGAGAAAATCAATTGCTAGTCACTAATGTGTGGTTTAAACTGGAGTGAACGACATG 60

QY 1114 AATCTCGCTGGAAACACTCGCATATGCGGAGTTAAGGATCTGCGAATACCCCGCAT 1173

Db 61 AATCTCGCTGGAAACACTCGCATATGCGGAGTTAAGGATCTGCGAATACCCCGCAT 120

QY 1174 CGCATCTGGAAGCGGACGCTGCTATGTACACAGTGGGATTTGACGGCACC 1233

Db 121 CGCATCTGGAAGCGGACGCTGCTATGTACACAGTGGGATTTGATGGCACC 180

QY 1234 TACCAGACGAACGTTGGTGGCGGAACAACGGCTCGTGTCTATACGTTCCGCGGGATC 1293

Db 181 TACCAGACGAACGTTGGTGGCGGAACAACGGCTCGTGTCTATACGTTCCGCGGGATC 240

QY 1294 TTCAAGTCGACGTCGAAGATCGACATCACTGGTTCCTTCGATGACCGGTCGGAG 1353

Db 241 TTCAAGTCGACGTCGAAGATCGACATCACTGGTTCCTTCGATGACCGGTCGGAG 300

QY 1354 ATGAAGTTCCGAGTTGGACCTACGACGATTCAGCTGGATTTACAATTACAAGATGAA 1413

Db 301 ATGAAGTTCCGAGTTGGACCTACGACGATTCAGCTGGATTTGATTTGGT----- 343

QY 1414 ACTGGCGGTGATATCAGCAGTTACGTGCTCAACGGCGAGTGGGAACCTACTGGTGGCC 1473

Db 344 -----TCAGTGTGCC 354

QY 1474 GGCAGAACGTAACGAGATCTATTCAACTGTGCGCGGAACCCCTATATAGACATCACCTTC 1533

Db 355 GGCAGAACGTAACGAGATCTATTCAACTGTGCGCGGAACCCCTATATAGACATCACCTTC 414

QY 1534 GCCATCATATCCGCGAGCAACACTGTACTATTCTTCAACCTGATCATCTTGTTGA 1593

Db 415 GCCATCATATCCGCGAGCAACACTGTACTATTCTTCAACCTGATCATCTTGTTGA 474

QY 1594 CTGATTGCCCTTCCATGGCTTGTCTGGATTCACCTCGCGGACGATTCGGTGAATAATTA 1653

Db 475 CTGATTGCCCTTCCATGGCTTGTCTGGATTCACCTCGCGGACGATTCGGTGAATAATTA 534

QY 1654 TCCTGGGTGTTACCAATCTTGTCTCTCGTACCGCTGTTCTGAATATGTTGCCGAGACA 1713

Db 535 TCCTGGGTGTTACCAATCTTGTCTCTCGTACCGCTGTTCTGAATATGTTGCCGAGACA 594

QY 1714 ATGCCGGCTACTTCCGATCGGTTGCCATTTGTGATAC 1750

Db 595 ATGCCGGCTACTTCCGATCGGTTGCCATTTGTGATAC 631

RESULT 5

AAZ24476

ID AAZ24476 standard; cDNA to mRNA; 3700 BP.

XX

XX AC AAZ24476;

XX

XX DT 17-FEB-2000 (first entry)

XX

XX DE H. virescens acetyl-choline receptor DNA from clone Hva7-1.

XX

XX KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening; neurotransmission; plant protection agent; conductance; AChR; ds.

XX

XX OS Heliothis virescens.

XX

XX FH Key Location/Qualifiers

XX CDS 335..1825

XX FT /*tag= a

XX FT /product= "acetyl-choline receptor"

XX

XX DE19819829-AL.

XX

XX PD 11-NOV-1999.

XX

XX PF 04-MAY-1998; 98DE-01019829.

XX

XX PR 04-MAY-1998; 98DE-01019829.

XX

XX PA (PARB) BAYER AG.

XX

XX PI Adamczewski M, Oellers N, Schulte T;

XX

XX DR WPI; 2000-014207/02.

XX

XX DR P-PSDB; AAY50815.

XX

XX PT New nucleic acid encoding a nicotinic acetylcholine receptor from insects, used to identify potential insecticides.

XX

XX PS Claim 1a; Page 14-17; 26pp; German.

XX

CC This invention describes a novel nucleic acid (NA) encoding a nicotinic acetyl-choline receptor (I) from insects which can be used as an insecticide. Inhibitors of (I) interfere with neurotransmission. (I) (also vectors containing it, its regulatory regions, and antibodies directed against (I)-encoded proteins) are used to screen for: (a) plant protection agents that alter conductance of AChR, potentially useful as insecticides, or (b) genes which encode polypeptides that are involved in formation of functionally related AChR in insects. (I) are also used to isolate and characterize the specified regulatory regions and for recombinant production of (II). This sequence encodes an acetyl-choline receptor isolated from *Heliothis virescens*

SQ Sequence 3700 BP; 893 A; 953 C; 944 G; 910 T; 0 U; 0 Other;

Query Match 22.2%; Score 512.8; DB 3; Length 3700;

Best Local Similarity 62.5%; Pred. No. 5.5e-133;

Matches 878; Conservative 0; Mismatches 502; Indels 24; Gaps 4;

QY 925 GGATATCATGAAAGAGACTGTACAGATCTTTGGATCTTATATACACTAGAACGT 984

Db 425 GGGTACCAAGAGCGGCTACTGACCACTATTGGACCACTACCACTGAGAGG 484

QY 985 CCCGTTCTCAATGAATCGGACCGCTTACAAATTAAGCTTTGGTTTAACTTAATGCAATT 1044

Db 485 CCCGTCGTCAACGAGAGCGACCGCTGACGCTCTCTTCGGCCTCAGCTCATGAGATC 544

QY 1045 ATCGATGTGGACGAGAAAATCAATTTGCTAGTCACTAATGTGTGGTTAAACTGGAGTGG 1104

Db 545 ATCGAGCTGGACGAGAGAACCCAGCTTTTAAACAAACATCTGGCTAAACTAGAGTGG 604

QY 1105 AACGATCAATCTCGGTGGAAACCTCCGACTATGCGGAGTTAAGGATCTGGAATA 1164

Db 605 AATGATATGAACCTTGAGTGGGAACACTTTCAGATTTTCGGCGGGTCAAAAGATTTAAAGAGTG 664
Qy 1165 CCGCGGATCGCATCTGGAGCGGACGCTGCTGATGTACAACTGCGGATGAGGATTT 1224
Db 665 CCACCCACAGACTATGGAACACAGAGCTCTTATGTACAAACGCGGAGGAGGTTG 724
Qy 1225 GAGCGCACTTACAGACGAACTGGTGGTGGGAAACAAAGCGCTGCTGTATAGTTCCG 1284
Db 725 GACAGCACTATCCAAAGACGTTGGTGGGAAACAAAGCGCTGCTGTCTAGTCCG 784
Qy 1285 CCGGGATCTTCAAGTCGAGCTGCAAGATCGACATCGACATCGAGTTCCCTTCGATGACAG 1344
Db 785 CCGGCACTTCAAGAGACACCTGCAAGATCGACATCGACATCGAGTTCCCTTCGAGACAA 844
Qy 1345 CGTTCGAGATGAAGTTCCGAGTTGACCTACAGGATTCAGCTGGATTTACAATTA 1404
Db 845 CGATCGAGATGAAGTTGGCAGCTGGAATTTATGATGTTATGATGTTGATGTTCAACTA 904
Qy 1405 CAAGATGAACCTGCGGTGATCAGCAGTTACGTGCTCAACGGCGAGTGGGAATCACTG 1464
Db 905 CAGATGAAGGGCGGAGATATAGCAGTTTGTACGATGGGATGGAGTTAATA 964
Qy 1465 GGTGTCCCGGCAACCTGAACGATCTTATACACTGCTGCCCGAACCTTATATAGAC 1524
Db 965 GGAGTCCCGGCAACGAGATCTACTACAACTGTTGTCGGAGCCATACATCGAC 1024
Qy 1525 ATCAGTTGCGCATCATCTCGCGGACGAACATGTTACTATTTCTTCAACTGATCATA 1584
Db 1025 ATCAGTTTGGGGTGGTATCCGGAGGAACCGCTCTACTACTTCTTCAATCTGATCGT 1084
Qy 1585 CTTGTGTACTGATTGCTTCATGCTTGTGCGGATTCAGCTGCCCGCAGATTCGGGT 1644
Db 1085 CCTGCGTGCATCGCTCATGCTCATGCTCTATGCGGTTCACTTGCCTCGAGCTCCGGA 1144
Qy 1645 GAAATATATCGTGGGTGTTTACCATCTTGTCTGCTGACCGTGTTCGAAATGATGTT 1704
Db 1145 GAAATGTTGCTTTAGGTGTGACGATATTACTGCTGTTGACGGTGTCTCTCAACATGGTG 1204
Qy 1705 GCGAGACAATGCGGCTACTTCCGATGGTGGCAATG-----TGGATACG 1752
Db 1205 GCGAGACGATGCCAGCAGCTGCGACGCGGTGCTTGTCTGCGGACCTACTTCACTGTC 1264
Qy 1753 ATCGTGTGTTTG--TGCTGGTGCATGGAATTCGGAATGAGTGCAGCCAGACACCGC 1810
Db 1265 ATCATGTTATGGTGGTCTCTCGCTGCTCCACCATACTGATCTCACTACACACAC 1324
Qy 1811 TGATCCTAGATTCGCGACACCGCTGTTGGACACATCTCTCCGAGCGAA-----GCAC 1866
Db 1325 CGGACGCGACACACTCAGAAATGAGTGTGATTCGTTGCTGCTTCTTATTTGGTGT 1384
Qy 1867 CAGATACTCTCGAGTTGAGCTGAAGAGCGCTGTCGAATCGCTGTCGCGCAACGTA 1926
Db 1385 CGTGGGTGCTGCGATGTCACGGCCGCTCGGCGACGACGCGCGCGCGCGCGTA 1444
Qy 1927 CTAGACATCGATGATGACTTCCGCGACAAATTCGCCCCATACGCGCGCGCGGAACACTG 1986
Db 1445 CTTCCGCGCGGACCTTGAGCTGCGGAGCGCTCTCCAGTCTCTCTACGAAAGCTG 1504
Qy 1987 CCACAAACCGCGCTTTCTATCGAGGTTTATGACAGGCGAGATGGGAGCATTTGG 2046
Db 1505 CTCGACATCGATGACGACTTCCGCGCACCGCGCGAGCGACGCGCAATGCTGCCGATAC 1564
Qy 2047 CCAATTCGACACCGGATCGGATCGGTGCTACCCATCATAGTGCATCAATCATCA 2106
Db 1565 TACAGGGGGGTGAGAGAAATGCGCGGGTGTGGCGCGCACAGTTGCTTC-----GGT 1618
Qy 2107 ACTGAATATGAATTAGGTTTAATCTTAAAGAAATTCGTTTATTAAGTACGATCAGTACGT 2166
Db 1619 GTCGACTACAGCTCTCCCTCATCTCTGAAGGAGATTAGAGTCAATCAGATCAGATCGCG 1678
Qy 2167 AAGATGACGAGTGCATGACATTCGCAATGATGGAATTTGCACTATGCTGCTGAC 2226
Db 1679 AAGGACGAGAGATCGGACATTTCCGCGCATGGAAGTTCCGCGCAATGCTGCTGGAC 1738

Qy 2227 AGACTGTGCTTATCATATTCACAAATGTCGCAATATTAGCCAAATAGTGTACTACTA 2286
Db 1739 AGACTGTGCTTATCATATTCACAAATGTCGCAATATTAGCCAAATAGTGTACTACTA 1798
Qy 2287 TCGGCACCAATATTTATGTTCTCG 2310
Db 1799 TCGCGCCACACATCATGTTGTGTCG 1822

RESULT 6
ABL13733
ID ABL13733 standard; cDNA; 936 BP.
XX ABL13733;
XX
DT 26-MAR-2002 (first entry)
XX
DB Drosophila melanogaster expressed polynucleotide SEQ ID NO 35681.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
OS WO200171042-A2.
XX PN
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB69630.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 35681; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB5773-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 936 BP; 238 A; 213 C; 239 G; 246 T; 0 U; 0 Other;

Query Match 20.7%; Score 479; DB 4; Length 936;
Best Local Similarity 73.5%; Pred. No. 8.3e-124;
Matches 611; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

Qy 913 GTTTCCTAGCAGGATATCATGAAAGAGACTGTTTACAGATCTTTTGGATCCTTATAT 972
Db 94 GCTTGTGGCTGTGACCCCATGAGAGGGCTACTCCAGCCCTTCTGGACAATACAAAC 153
Qy 973 ACACCTAGAACCTCCCGTTTCTCAATGAATCGGACCGCTTACAAATTAAGCTTTGGTTAACT 1032
Db 154 ASCCTGGAGCGTCCCGTGTGTCATGAATCCGATCCATTCGAACTGAGCTTCGAGCTAACAA 213
Qy 1033 TTAATGCAAAATTATCGATGTGACGAGAAATCAATTCCTAGTACTAATGTGTGTTA 1092

Db 214 CTCATGCAGATTATCGATGTGCGAAGAAAGATCAACTGCTTATACGATATTTGGCTC 273
 QY 1093 AAACCTGGAGTGGACGACATGAATCTCCGCTGGGAACACCTCCGACTATGGCGGAGTTAAG 1152
 Db 274 AAATTTGGAAATGGAACGATGAATCTTCGATGGAATTCGAGTGAATTCGGTGGTGGCG 333
 QY 1153 GATCTGCCAATACCGCCGATCGCATCTGGAGCGGACGCTGCTGATGTACAAAGTGG 1212
 Db 334 GATCTGCCAATTCGCCACATCGCTATGGAACCGGATGACTGATGATACAAAGTGGCC 393
 QY 1213 GATGAGGAAATTGACGGCACTTACGACGACGACGCTGTGGTGGCGAAACACGCTCGTGT 1272
 Db 394 GACGAGGCTTCGATGGAACGCTACGCCACAAATGTGGTGGTTCGCAATATGGGAGCTGT 453
 QY 1273 CTATACGTTCCCGCGGGATCTTCAAGTCGACGTGCAAGATCGACATCACTGGTTCGCC 1332
 Db 454 CTGTACGTACCGCCAGGTATATTTAAGTCAACGTTAAGATCGACATTTACGTGGTTCCA 513
 QY 1333 TTCGATGACCAAGCGGTGCGAGATGAAGTTCCGCGAGTTGGACCTACGACGATTCAGCTG 1392
 Db 514 TTCGACGATCAGAGATGTGAATGAATTTGGTTCGTGGACCTACGATGGGTTTCAGTTG 573
 QY 1393 GATTACAAATTACAGATGAACCTGGCGGTGATATCAGCAGTTACGTCTCAACGGCGAG 1452
 Db 574 GACCTGCAGTTGCAGGACGAAGCTGGTGGCGACATTTCTAGCTTTATAACCAATGGCGAA 633
 QY 1453 TGGGAACCTACTGGGTGTGCGCGCAAGCTGTAACAGATCTATTACAACTGCTGCGCGGAA 1512
 Db 634 TGGGACTTGTGTAGGTGTGCGCGTAAACGAATGAATCTACTATAATTTGTCGCCAGAA 693
 QY 1513 CCCTATATAGACATCACTTCGCCATCATCATCGCGACGACACACTGTACTATTCTTC 1572
 Db 694 CCTTATATTGACATAACATTCGCCATTTTGAATAGCGCGAAAGCTTGTACTATTTTTC 753
 QY 1573 AACCTGATCATACCTGTGTACTGATGATTCCTCCATGCGCTTGTCTCGGATTCACCGCTCGG 1632
 Db 754 AATCTGATTTGTGCGTGCCTACTGATCGCTCCATGSCACTGCTAGGTTTACACTGCCA 813
 QY 1633 CCAGATTCCGGTGAATAATTTACTGCTGGGTGTTACCATCTTGCTCTGCTGACCGTGT 1692
 Db 814 CCAGATTCTGGTGAAGAAGCTTTCGCTTGGAGTTACAAATCTATTATCGCTTACAGTCTTC 873
 QY 1693 CTGAATATGTTGCCGAGACAAATGCCGGCTACTTCCGATCGGTGCGCATG 1743
 Db 874 CTCACATGTTGCCGGAACAAATGCCGCGACCTCCGATCGGTACCGCTG 924

RESULT 7

ABL07231

XX ABL07231 standard; cDNA; 1540 BP.

XX AC

XX ABL07231;

XX

DT 26-MAR-2002 (first entry)

XX

XX

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 16175.

XX

XX

XX Drosophila; developmental biology; cell signalling; insecticide;

XX

XX pharmaceutical; gene; ss.

XX

XX Drosophila melanogaster.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PI

XX

XX

DR

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Venter JC, Adams M, Li PWD, Myers EW;

WPI; 2001-656860/75.

P-PSDB; ABB63128.

New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signaling and cell-cell

interactions.

Claim 1; SEQ ID NO 16175; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent

capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of

insecticides, therapeutics and pharmaceutical drugs. The invention

discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA

sequences (ABU01840-ABU16175) and the encoded proteins (ABBS57737-

ABH2072). The sequence data for this patent did not form part of the

printed specification, but was obtained in electronic format directly

from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1540 BP; 412 A; 390 C; 376 G; 362 T; 0 U; 0 Other;

Query Match 17.8%; Score 412.2; DB 4; Length 1540;

Best Local Similarity 69.2%; Pred. No. 6.8e-105;

Matches 578; Conservative 0; Mismatches 253; Indels 4; Gaps 1;

909 TAAAGTTTGCTAGCAGGATATCATGAAAGAGACTGTTACAGATCTTTTGGATCTTAA 968

106 TAAAGAAAGCTGTCAAGGACCTCATGAAAGCGCTCTGTAACCATCTGCTGCCACTA 165

969 TAATACACTAGAACGTCCTCTCAATGAATCGGACCCGTACAAATAGACTTGGTTT 1028

166 CANTACGCTGGAGCGACCCGTGCCAATGAATCGGAGCCCTGGAGGTTAAGTTTCGACT 225

1029 AACTTAAATGCAAAATTATCGATGTGACGAGAAATCAATTCGTAGTCACTAATGTGTG 1088

226 GACGCTGCAGCAGATCATCGACGTGGATGAAAGAAATCAGATTCAGACCAATGCGTG 285

1089 GTTAAAACTGGAGTGAACGACATGAATCTCCGCTGGAAACACCTCCGACTATGCGGAGT 1148

286 GTTAAATTTGGAGTGAACGACTACAATCTCGCTGGAAATGAAACGGAATACGCGGGT 345

1149 TAAGGATCTCGGAATATACCGCGCATCGCATCTCGAAGCGGAGCTGTGTGATGTACACAG 1208

346 CAAGGATCTACGATCAGCCCAACGAGCTGTGGAAGCCGCGCTCAATGATACACAG 405

1209 TCGGATGAGGGATTTGACGGCACCTACAGACGAACTGGTGGTGGCGAACAACGGCTC 1268

406 CGCGATGAGGGATTCGATGGCAGCTATCACACCAATTTGTTGTTCAAAATACCGGAG 465

1269 GTGCTATACGTTCCGCGGGGATCTTCAAGTCGACGTGCAAGATCGACATCACTGGTGT 1328

466 TTGCTGTACGTCGCCCTCGTATCTTCAAGAGCACATCAAGATAGACATCACTGGTGT 525

1329 CCCCTTCGATGACCGCGGTGGAGATGAAGTTCCGAGTTGGACCTTACGACGATTCGA 1388

526 CCCATTTGATGACCAACATTTGGAATGAAATTCGTTAGTGGACTTACGATGGAATCA 585

1389 GCTGGATTTACAAATTACAGATGAACCTGGCGGTGATATCAGCAGTTACGTGTCAACGG 1448

586 GTTGGATTTGGTTTGAATTCGGAATGAGGGGATCTTTCCGATTTTCAACAAATGG 645

1449 CGAGTGGGACTACTGGGT----GTGCGCGGCAAGTAAACAGATCTATTACAATGCT 1504

646 CGAGTGGTACTTGTGTTGTGTTACCATTCGCGGGAAGAAAGATACGATAGTCTACGCTGT 705

1505 GCCCGGAACCCCTATATAGACATCACTTCGCGCATCATCATCGCGGAGCAACACTGTACT 1564

706 GCCGAGAACCATATGTCGATATCACCTTACTATACAAATTCGTCGCCGATCATATATT 765

QY 1091 TAAACCTGGAGTGAACACATGAATCTCCGCTGGAAACACCTCCGACTATGGGGAGTTA 1150
 Db 233 TGCATAATGCTTGGACAGATCACTATTATCAGTGGATGTGTAGATATATCCAGGGGTGA 292
 QY 1151 AGGATCTGCGAATACCCCGCATCGCATCTGGAGCCGAGCGTGTCTGATGTACAAAGTG 1210
 Db 293 AGACTGTTGGTTCCTCCAGATGGCCAGATTTGGAAACCCAGACATTTCTTCTATACAGTG 352
 QY 1211 CGGATGAGGATTTGACGGCACCTACAGACGAAACGTTGGTGGTGGCGGAAACAAACGGCTCGT 1270
 Db 353 CTGATGAGCGCTTGGACGCCACATTCACACTAAACGTTGGTGGTAAATCTTCTGGGCAAT 412
 QY 1271 GTCTATACGTTCCGCGGGGATCTCAAGTCGACGTCGAGATCGACATCACGTTGGTTC 1330
 Db 413 GCCAGTACCTGCTCCAGGATATTCAGAGTCTCTGTACATCGATGTACGTTGGTTC 472
 QY 1331 CCTTCGATACCGAGCGGTGGAGATGAAGTTGGCAGTTGGACCTACGACGATTCGACG 1390
 Db 473 CCTTTGATGTGACGACCTGCAAACTGAAGTTGGGTCTGTCTTACGGAGCTGTGCTCT 532
 QY 1391 TGGATTTACAAATACAGATGAAGTGGCGGTGATATCAGCAGTTACGTTGCTCAACGCG 1450
 Db 533 TGGATCTGCAGATCAGAGGCA-----GATATCAGTGGCTATATCCCAATGGAG 583
 QY 1451 AGTGGGAATCTACTGGGTGTGCCGCAAAACGTAACGAGATCTATTACAACTGCTGCCCGG 1510
 Db 584 AATGGGACCTAGTGGATATCCCGGACGAGAGTGAAGGTTCTATCAGTGTGCAAG 643
 QY 1511 AACCTATATAGACATCACTTCGCGCATCATCTCGCGGACGACACTGTACTATTCTT 1570
 Db 644 AGCCCTACCCGATGTACCTTTCAGTGAACATGCGCGCGGACGCTCTACTATGGCC 703
 QY 1571 TCACCTGATCATACCTTGTGTGATGTGCTCCATGCGCTTGTGCTGGATTCACCTGTC 1630
 Db 704 TCACCTGCTGATCCCTGTGTGCTCATCTCCGCGCTGCTGCTGTGTCTGCTTTC 763
 QY 1631 CGCCAGATTCGGGTGAATAATATCGCTGGGTGTATCCATCTGTCTCGCTGACCGCTGT 1690
 Db 764 CTGACGATTCGGGAGAGATTTCCCTGGGATAACAGTCTTACTCTCTTACCGCTCT 823
 QY 1691 TTCTGAATATGGTTCGCGAGACAATGCGCGTACTTCCGATGCGGTGCGCATTTG 1743
 Db 824 TCATGCTGCTGCTGCTGAGATCATGCCGCAACATCCGATTCGTTACCATTTG 876

RESULT 11
 AAC90380
 ID AAC90380 standard; cDNA; 1509 BP.
 XX AC
 AC AAC90380;

XX 14-MAR-2001 (first entry)

XX Wild-type human alpha7 ligand gated ion channel coding sequence.

XX Human; alpha7 nicotinic acetylcholine gated ion channel;
 KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.

XX Homo sapiens.

XX WO200073431-A2.

XX 07-DEC-2000.

XX 25-MAY-2000; 2000WO-US011862.

XX 27-MAY-1999; 99US-0136174P.

XX (PHAA) PHARMACIA & UPJOHN CO.

XX Groppi VE, Wolfe ML, Berkenpas MB;

XX

DR NFI; 2001-061524/07.
 DR P-PSDB; AAB50012.
 XX
 PT Special cell culture medium for treating cells and for inducing mammalian
 PT cell lines to conduct calcium ions, comprising specified concentrations
 PT of ions of sodium, calcium and potassium at specified pH.
 XX
 PS Example 5; Page 60-61; 77pp; English.
 XX
 CC The present sequence is the coding sequence for wild-type human alpha7
 CC nicotinic acetylcholine gated ion channel. The human alpha7 ion channel
 CC was used in the construction of an alpha7/5-hydroxytryptamine (5-HT3)
 CC chimeric ligand gated ion channel (see AAC90382 and AAB50014). The
 CC alpha7/5-HT3 chimeric ion channel can be expressed by recombinant cells
 CC in the present invention, resulting in preferential calcium ion
 CC conductance by the cells
 XX
 SQ Sequence 1509 BP; 297 A; 451 C; 429 G; 332 T; 0 U; 0 Other;

Query Match 12.8%; Score 296.2; DB 4; Length 1509;
 Best Local Similarity 60.7%; Pred. No. 2.9e-72;
 Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;

QY 911 AAGTTTCCTAGCAGGATATCATGAAAGAGACGCTTTACAGATCTTTTGGATCCTATA 970
 Db 53 ACGTGTCCCTGCAAGCGAGTTCCAGAGGAAGCTTTTACAAGAGCTGGTCAAGAACTACA 112
 QY 971 ATACACTAGAACGTCCTGCTTCAATGAATCGGACCGCTTACAATTAAGCTTTGGTTTAA 1030
 Db 113 ATCCCTTGGAGAGCCCGTGGCCAAATGACTCGCAACCACTCACCGTCTACTTCTCCCTGA 172
 QY 1031 CTTTAAATGCAAAATPATCGATGTGGACGAGAAATCAATTTCTAGTACATTAATGTGGT 1090
 Db 173 GCCTCTTCGAGATCATGCACTGTGATGAGAAGAACCAAGTTTAAACCAACATTTGGC 232
 QY 1091 TAAACTGGAGTGAAGACATGATGATCTCCGCTGGAACACCTCCGACTATGCGGAGTTA 1150
 Db 233 TGCAAATGTCTTGGACAGATCACTATTACAGTGGAAATGTGTCAGATATCCAGGGTGA 292
 QY 1151 AGGATCTGCGAATAACCCCGCATCGCATCTGGAAGCCGAGCGTCTGATGTACAACTG 1210
 Db 293 AGACTGTGTTTCCAGATGGCCAGATTTGGAACCAAGACATTTCTTCTATAACAGTG 352
 QY 1211 CGGATGAGGATTTGACGGCACCTTACAGACGAAAGTGTGTGTGCGGAAACAAACGCTCGT 1270
 Db 353 CTGATGAGCGCTTTGACGCCACATTCACACTAAACGTTGTGTAATTTCTTCGGCAAT 412
 QY 1271 GTCTATACGTTCCGCGGGATCTTCAAGTCGAGTCAAGATCGACATCGACATCGTGTTC 1330
 Db 413 GCCAGTACCTGCTCCAGGCATATTCAGAGTTCTCTGCTCATCATGATGACGTTGTTTC 472
 QY 1331 CTTTCGATGACGAGCGGTGCGAGATGAAGTTTCGGCAGTTGGACCTTACGACGATTCACG 1390
 Db 473 CTTTGTGATGCGACACTGCAAACTGAAGTTTGGTCTGCTTCTACGAGGCTGGTCT 532
 QY 1391 TGGATTTACAAATTAAGATGAATACTGGCGGTGATATCAGCAGTTACGTGTCAACGGG 1450
 Db 533 TGGATCTGAGATGCGAGGCA-----GATATCAGTGGCTATATCCCAATGGAG 583
 QY 1451 AGTGGGACTACTGGGTGTGCCGCAAAACGTAACGAGATCTATTACAACTGCTGCCCGG 1510
 Db 584 AATGGGACCTAGTGGAAATCCCGGCAAGAGAGTGAAGGTTCTATGAGTGTGTCGAAG 643
 QY 1511 AACCTATATAGACATCACTTCCCATCATCATCCGCGGACGACACTGTACTATTCTT 1570
 Db 644 AGCCCTACCCGATGTACCTTTCAGTGAACATGCGCGCGGACGCTCTACTATGGCC 703
 QY 1571 TCACCTGATCATACCTTGTGTGATGTGCTCCATGCGCTTGTGCTGGATTCACCTGTC 1630
 Db 704 TCACCTGCTGATCCCTGTGTGCTCATCTCCGCGCTGCTGCTGTGTCTGCTTTC 763
 QY 1631 CGCCAGATTCGGGTGAATAATATCGCTGGGTGTATCCATCTGTCTCGCTGACCGCTGT 1690

Db 764 CTGCAGATTCGGGGAGAGATTTCCCTGGGGATAACAGCTTACTCTCTTACCGTCT 823
 QY 1691 TTCTGAATATGTTGGCCGAGACAAATGCGGCTACTTCCGATGCGTGCATTG 1743
 Db 824 TCATGCTGCTGCTGGCTGAGATCATGCCCCGAACATCGGATTCGGTACCAATTG 876
 RESULT 12
 AAV12197
 ID AAV12197 standard; cDNA; 1876 BP.
 XX AAV12197;
 AC
 DT 14-MAY-1998 (first entry)
 XX
 DE Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.
 KW Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
 KW brain tissue; screening; NACHR; antibody; ds.
 XX Homo sapiens.
 OS
 PH Key Location/Qualifiers
 FT CDS 73..1581
 FT /*tag= a
 FT /product= "neuronal nicotinic acetylcholine receptor
 FT alpha-7 subunit"
 XX
 PN WO9420617-A2.
 XX
 PD 15-SEP-1994.
 XX
 PF 08-MAR-1994; 94WO-US002447.
 XX
 PR 08-MAR-1993; 93US-00028031.
 XX
 PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
 XX
 PI Elliott KJ, Ellis SB, Harpold MM;
 XX
 DR WPI; 1994-303024/37.
 DR P-PSDB; AAW44153.
 XX
 CC Human neuronal nicotinic acetylcholine receptor subunits and DNA - also
 CC transformed cells useful for screening cpds. which modulate activity of
 CC the receptor.
 PS Claim 8; Page 78-79; 99pp; English.
 XX
 CC The present sequence encodes a human neuronal nicotinic acetylcholine
 CC receptor (NACHR) subunit. The cells expressing the alpha and/or beta
 CC NACHR subunits may be used in a method of screening compounds to identify
 CC any which modulate the activity of human neuronal NACHR. Subunit specific
 CC antibodies may be used to monitor the distribution and expression density
 CC of various subunits in normal vs diseased brain tissues. Testing of
 CC single receptor subunits or specific receptor subunit combinations with a
 CC variety of potential agonists or antagonists provides information with
 CC respect to the function and activity of the individual subunits and
 CC should lead to the identification and design of compounds that are
 CC capable of very specific interaction with one or more receptor subtypes.
 CC The resulting drugs should exhibit fewer unwanted side effects than drugs
 CC identified e.g. screening with cells that express a variety of subtypes
 CC
 SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;
 Query Match 12.8%; Score 296.2; DB 2; Length 1876;
 Best Local Similarity 60.7%; Pred. No. 3.3e-72;
 Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;
 QY 911 AAGTTTGCCTAGCAGATATCATGAAGAAGACATGTTACAGATCTTTTGGATCCTTATA 970
 Db 125 ACCTGTCCTGCAAGCGAGTTCAGAGGAAGCTTTACAGAGGAGTGGTCAAGAACTACA 184

QY 971 ATACACTAGAACGTCCCGTCTCAATGAATCGGACCCGTTACAATTAAGCTTTGGTTAA 1030
 Db 185 ATCCCTTTGGAGAGCCCGTGGCCAAATGACTCGCAACCACTCACGCTCTACTTCTCCCTGA 244
 QY 1031 CTTTAAATGCAAAATATCGATGTGACGAGAAAATCAATTCCTAGTCACTAATGTGTGGT 1090
 Db 245 GCCTCTCTGCAGATCATGGACGTGGATGAGAAGAACCAAGTTTAAACCAACAATTTGGC 304
 QY 1091 TAAAACTGGAGTGGACGACATGAATCTCCGCTGGAAACACCTCCGACTATGGCGAGTTA 1150
 Db 305 TGCAAAATGTCTTGGACAGATCACTATTTTACAGTGGAAATGTGTGAGATATCCAGGGTGA 364
 QY 1151 AGGATCTGCGAATACCCCGCATCGATCTGGAAGCCGAGCTGCTGATGATPACACAGTG 1210
 Db 365 AGACTGTTCCTTCCAGATGGCCAGATTTGGAACCAAGACATTTCTTCTCTATAACAGTG 424
 QY 1211 CGGATGAGGGAATTTGACGGCACCTPACCAAGACGAGTGGTGGTGGGGAACAACCGCTCGT 1270
 Db 425 CTGATGAGCGCTTTGACGCCACATTCACACTAAGTGTGGTGGAAATCTTCTGGGCATT 484
 QY 1271 GTCTATACGTTCGCGCGGGGATCTTCAAGTCGAGCTGCAAGATCGACATCACGTGGTTC 1330
 Db 485 GCCAGTACCTGCCTCCAGGCATATTCAAGAGTTCTCTCTACATCGATGACGTGGTTTC 544
 QY 1331 CCTTCGATGACCGGTGCGAGATGAAGTTCCGCACTTGGACCTACGACGAGATTCGAGC 1390
 Db 545 CCTTGTATGTGCAGCACTGCAAACTGAAGTTTGGGTCTCTGTACGAGGCTGGTCTCT 604
 QY 1391 TGGATTTACAATTAAGATGAAACTGGCGGTGATATACAGCAGTTACGTCTCAACGGCG 1450
 Db 605 TGGATCTGCAGATGCAGGAGCA-----GATATCAGTGGCTATATCCCAATGGAG 655
 QY 1451 AGTGGGAATCTACTGGGTGTGCCCGCAAAAGTAAACGAGATCTATTACAATCTGTGCCCGG 1510
 Db 656 AATGGGACCTAGTGGGAATCCCCGGCAAGAGGAGTGAAGGTTCTATGAGTGTGCAAG 715
 QY 1511 AACCTATATAGACATCACCTTCGCCATCATATCCGCCCTCGCCCTGCTGGTGTCTCTGTTCT 1570
 Db 716 AGCCCTACCCCGATGTACCTTCACAGTGACCATGCGCCGACGACGCTCTACTATGGCC 775
 QY 1571 TCAACCTGATCATACCTTGTGTACTGATTCCTCCATGGCTTGTCTCGGATTCACCCCTGC 1630
 Db 776 TCAACCTGTGATCCCTGTGTGCTCATCTCCGCCCTCGCCCTGCTGGTGTCTCTGTTCT 835
 QY 1631 CGCCAGATTCGGTGAAAAATTAATCGCTGGGTGTGTACCATCTTGTCTCTCGCTGACCCGTG 1690
 Db 836 CTGCAGATTCGGGGAGAGATTTCCTGGGGATAACAGTCTTACTCTCTCTTACCGTCT 895
 QY 1691 TTCTGAATATGTTGGCGAGACAATGCGGGCTACTTCCGATGGGTGGCAATTG 1743
 Db 896 TCATGCTGCTGCTGGCTGAGATCATGCCGCAACATCCGATTCGTACCAATTG 948
 RESULT 13
 AAT48239
 ID AAT48239 standard; DNA; 1876 BP.
 XX
 AC AAT48239;
 XX
 DT 09-APR-1997 (first entry)
 XX
 DE Neuronal nicotinic acetylcholine receptor alpha-7 subunit DNA.
 XX
 KW Neuronal nicotinic acetylcholine receptor; nAChR; neurotransmitter;
 KW ligand-gated receptor; ds.
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT CDS 73..1581
 FT /*tag= a
 XX
 PN WO9641876-A1.

XX PD 27-DEC-1996.
 XX PF 07-JUN-1996; 96WO-US009775.
 XX PR 07-JUN-1995; 95US-00484722.
 XX PA (SIBI-) SIBIA NEUROSCIENCES INC.
 XX PI Elliott KJ, Harpold MW;
 XX WPI; 1997-065463/06.
 XX P-PSDB; AAM09025.
 XX Nucleic acids encoding nicotinic acetylcholine receptor sub-units - used
 PT in screening to determine the effect of drugs on the receptor.
 XX PS Disclosure; Page 71-73; 108pp; English.
 XX CC A DNA sequence (AAM48239) codes for the alpha-7 subunit (AAM09025) of the
 CC human neuronal nicotinic acetylcholine receptor (nAChR). Host cells, esp.
 CC mammalian cells or amphibian oocytes, carrying alpha-7 nucleic acids,
 CC opt. in combination with other alpha and/or beta subunit nucleic acids
 CC (see also AAT48232-38, AAT48240-41), express recombinant nAChR subunits
 CC useful for identifying cpds. that modulate the activity of human nAChRs
 XX SQ Sequence 1876 BP; 369 A; 553 C; 531 G; 423 T; 0 U; 0 Other;
 Query Match 12.8%; Score 296.2; DB 2; Length 1876;
 Best Local Similarity 60.7%; Pred. No. 3.3e-72;
 Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;
 QY 911 AAGTTTCCCTAGCAGGATATCATGAAGAAGACACTTTTACAGATCTTTTGGATCCTTATA 970
 DB 125 ACGTGTCTCCCTCAAGGCGAGTTCAGAGGAAGCTTTTCAAGGAGCTGTCTAAGAACTACA 184
 QY 971 ATACACTAGAACGTCCTCTCAATGAATCGGACCCGCTTACAAATTAAGCTTTGGTTTAA 1030
 DB 185 ATCCCTTGGAGAGGCCGTGGCCCAATGACTCGCAACCACTCACCGTCTACITTCCTCTGA 244
 QY 1031 CTTTAAATGCAATTTATGATGTGACGAGAGAAATCAATTCCTAGTCTACTAATGTGTGGT 1090
 DB 245 GCCTCTCTGCAGATCATGACGTGTGATGAGAAGAACCAAGTTTAAACCAACCAATTTGGC 304
 QY 1091 TAAACTGGAGTGAACACATGAATCTCCGCTGGAACACCTCGCACTATGCGGAGTTA 1150
 DB 305 TGCATAATGCTTGGACAGATCACTATTTACAGTGAATGTGTCAAGATATCCAGGGTGA 364
 QY 1151 AGGATCTCGGAATACCGCCGATCGCATCTGGAAGCCGAGCGTGTGATGTACAAAGTG 1210
 DB 365 AGACTGTTCGTTTCCAGATGGCCAGATTTGGAAACCAAGACATCTTCTCTATAACAGTG 424
 QY 1211 CGGATGAGGGAATTTGACGGACCTTACGAGAACGCTGTGTGGTGGGAAACACGCTCGT 1270
 DB 425 CTGATGAGCGCTTTGACGCCACATTTCCACACTAAGCTGTGTGTAATCTTCTGGGCATT 484
 QY 1271 GTCTATAGTTCGCGCGGGATCTTCAAGTGCAGTGCAGATCGACATCAAGTGTGTTCC 1330
 DB 485 GCCAGTACTCGCTCCAGGCAATTTCAAGATTTCTGCTACATCATGATGATGCTGTTC 544
 QY 1331 CTTTCGATGACAGCGGTGCGAGATGAAGTTTCGGCAGTTGGACCTTACGACGATTTCCAGC 1390
 DB 545 CTTTGTATGTGACGACTGCAAACTGAAGTTTGGGTCTGTGTTTACGGAGCGTGGTCT 604
 QY 1391 TGGATTACAAATTAAGATGAACTGCGGTGTATCATGACGATTTACGTGTCTCAAGCGG 1450
 DB 605 TGGATCTGAGATGACGAGGCA-----GATATCAGTGGCTATATCCCAATGAG 655
 QY 1451 AGTGGGAACACTAGTGGTGTGCGCGCAACACGTAACGAGATCTATTTACAACTGCTCCCGG 1510
 DB 656 AATGGGACCTAGTGGGAATCCCGGCAAGAGAGGTGAAGGTTCTATGAGTGTCTGCAAG 715
 QY 1511 AACCTATATAGACATCACCTTTCGCCCATCATCCCGCGAGAACACTGTACTATTCTT 1570

DB 716 AGCCCTACCCCGATGTCACTTCCACAGTGACCATGCGCCGACGACCTCTACTATGGCC 775
 QY 1571 TCAACCTGATCATACCTTGTGTACTGTATGCTTCCATGCGCTTGTCTCGGATTCACCCCTGC 1630
 DB 776 TCAACCTGCTGATCCCTCTGTGTCTATCTCCGCCCTCGCCCTGCTGTGTCTCTGCTTC 835
 QY 1631 CGCCAGATTCGGGTGAAAAATATCGCTGGGTGTACCAATCTTGTCTCGCTGACCGTGT 1690
 DB 836 CTGCAGATTCGGGGGAGAGATTTCCTCGGGATAACAGTCTTACTCTCTTACCGTCT 895
 QY 1691 TTCTGAATATGTTGCGGAGACAAATGCGGCTACTTCCGATCGGTGCCATTG 1743
 DB 896 TCATGCTGCTGTGCTGATCATGATCGCCGCAACATCCGATTCGTTACCATTTG 948
 RESULT 14
 ABS54875
 ID ABS54875 standard; cDNA; 1876 BP.
 XX AC ABS54875;
 XX DT 06-DEC-2002 (first entry)
 XX DE Human neuronal nicotinic acetylcholine receptor alpha 7 subunit cDNA.
 XX KW Human; neuronal nicotinic acetylcholine receptor; nNACHR; gene; ss;
 XX KW ion flux; alpha 7 subunit.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 73..1581
 FT /*tag= a
 FT /product= "Human nNACHR alpha 7 subunit"
 XX US6440681-B1.
 XX 27-AUG-2002.
 XX 07-JUN-1995; 95US-00487596.
 XX 03-APR-1990; 90US-00504455.
 XX 30-NOV-1992; 92US-00938154.
 XX 08-MAR-1993; 93US-00028031.
 XX 08-NOV-1993; 93US-00149503.
 XX (MERI) MERCK & CO INC.
 XX Elliott KJ, Ellis SB, Harpold MW;
 XX WPI; 2002-711528/77.
 XX P-PSDB; ABS70492.
 XX Identifying antagonists or agonists of human neuronal nicotinic
 PT acetylcholine receptors, by contacting recombinant cells with test
 PT compound, and measuring ion flux of cells or binding of compound to
 PT nNACHR.
 XX Claim 101; Col 57-60; 56pp; English.
 XX The invention relates to a method for identifying compounds that are
 CC antagonists or agonists of human neuronal nicotinic acetylcholine
 CC receptors (nNACHRs), by contacting recombinant cells with a test compound
 CC and measuring ion flux, the electrophysiological response of the cells or
 CC binding of the test compound to the nNACHR. The recombinant cells are
 CC produced by transfection with a nucleic acid encoding at least one human
 CC nNACHR (alpha or beta) subunit, such that the cells express an nNACHR
 CC comprising one human subunit encoded by the transfected nucleic acid.
 CC This sequence represents cDNA encoding the alpha 7 subunit of the human
 CC nNACHR polypeptide
 XX Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Query Match 12.8%; Score 296.2; DB 6; Length 1876;
Best Local Similarity 60.7%; Pred. No. 3.3e-72;
Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;

QY 911 AAGTTTGCTAGCAGGATATCATGAAAGAGAGACTGTTACAGCATCTTTGGATCCTTATA 970
Db 125 ACGTGTCCCTGCAAGCGAGTTCAGAGGAGCTTTTACAGGAGCTGCTCAAGACTACA 184
QY 971 ATACACTAGAACGTCCTGCTCTCATGATCGGACCCGTTCAATTAAGCTTTGGTTTAA 1030
Db 185 ATCCCTTGGAGAGGCCGTGGCCATGACTCGCAACCACTCACCGTCTACTTCTCCCTGA 244
QY 1031 CTTTAATGCAATATTCGATGGGACGAGAAAAATCAATTGCTAGTCAATGTGTGT 1090
Db 245 GCCTCTCCAGATCATGACGTGGATGAGAGAACCAAGTTTAAACCAACATTTGGC 304
QY 1091 TAAACTGGAGTGGAAACGACATGATCTCCGCTGGACACACCTCCGACTATGCGGAGTTA 1150
Db 305 TSCAAATGCTTTGGACAGATCACTATTACAGTGGATGTGTCAGATATCCAGGGTGA 364
QY 1151 AGGATCTCGGAATACCGCGCATCGCATCTCGAAGCCGAGCTGCTGATGTACAACTGTG 1210
Db 365 AGACTGTTCGTTCCAGATGCCAGATTTGGAAACCAAGATTTCTCTATAACAGTG 424
QY 1211 CGGATGAGGGATTTGACGGCACCTACCAAGACGAGTGGTGGTGGGAAACACGCTCGT 1270
Db 425 CTGATGAGCGTTTGGACGACATTCACATTAACGTTGGTGAATTTCTTGGGCAAT 484
QY 1271 GTCTATAGTTCGCCCGGGATCTTCAAGTCGAGTCGACATCGACATCACTGCTTCC 1330
Db 485 GCCAGTACCTGCTCCAGGCATATCAAGATTTCTGCTACATGATGACGTGTTTC 544
QY 1331 CTTTCGATGACCGGTGGAGATGAAGTTCGCGAGTGTGAACCTAGACGGATTCAGC 1390
Db 545 CCITTGATGCGACACTGCAAACTGAAGTTGGTCTCTTACGGAGCTGGTCT 604
QY 1391 TGGATTTCAATTAACAAGATGAACCTGCGGTGATATCAGCAGTTACGTCTCAACGCG 1450
Db 605 TGGATCTGAGATCGCAGGAGCA-----GATATGATGGCTATATCCCAATGAG 655
QY 1451 AGTGGGAATCTAGGTGGTGGCCGGGAAACGTAAACGAGATCTATTACAACTGCTGCCGG 1510
Db 656 AATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGTTCTATGAGTGTGCAAG 715
QY 1511 AACCTATATAGACATCACTTCGCATCATCTCCGCGCATCATCCGCGGACGACATGCTATTCT 1570
Db 716 AGCCCTACCCGATGTCACCTTCACAGTGACCATGCGCCGACGACGCTCTACTATGGCC 775
QY 1571 TCAACCTGATCATACCTTGTGTACTGATTCCTCCATGGCTTGTCTCGGATTCACCTGC 1630
Db 776 TCAACCTGCTGATCCCTGTGTCTCATCTCCGCTTGTCTGCTGCTGCTGCTGCTGCTGCT 835
QY 1631 CGCCAGATTCGGTGAAAATATTCGCTGGGTGTTACCATCTTCTGCTCGCTGACCGGTGT 1690
Db 836 CTGAGATTCGGGGAGAGATTTCCCTGGGGATAACAGTCTTACTCTCTTTACCGTCT 895
QY 1691 TTCTGAATATGTTGCGACACATCCCGCTACTTCCGATGGGTGGCCATTG 1743
Db 896 TCACTGCTGCTGGGTGAGATCATGCCCAACATCCGATTCGATACCATTTG 948

RESULT 15

ABV73248
ID ABV73248 standard; cDNA; 1876 BP.
XX AC ABV73248;
XX AC ABV73248;
DT 22-JAN-2003 (first entry)
TX Human neuronal NACHR alpha7 subunit encoding cDNA.
DE Human; neuronal; nicotinic acetylcholine receptor; NACHR; drug screening;
XX KW

immunohistochemistry; NACHR alpha7 subunit; gene; ss.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT CDS 73..1581
XX FT /*tag= a
XX FT /product= "NACHR alpha7 subunit"
XX FT /note= "neuronal nicotinic acetylcholine receptor"
XX PN WO200259266-A2.
XX PD 01-AUG-2002.
XX PF 29-OCT-2001; 2001WO-US050985.
XX PR 01-NOV-2000; 2000US-00703951.
XX PA (MERI) MERCK & CO INC.
XX PI Gillespie A, Claeps BO, Chavez-Noriega LE, Siegel R, Elliott KJ;
XX WPI; 2002-698532/75.
XX DR P-PSDB; ABB82435.
XX CC Cell comprising nucleic acids encoding human alpha and beta subunits of neuronal nicotinic acetylcholine receptors, useful for in vitro screening of a drug substance in a test system specific for humans.
XX PS Example; Page 129-130; 143pp; English.
XX CC The invention relates to a suitable host cell transfected with an isolated nucleic acid molecule comprising a sequence of nucleotides or ribonucleotides that encodes at least one alpha or beta subunit of a human neuronal nicotinic acetylcholine receptor (NACHR). The compositions and methods of the present invention, which provide a means to prepare synthetic or recombinant receptors and receptor subunits that are substantially free of contamination from many other receptor proteins, are useful for observing the effect of a drug substance on a particular subtype to perform in vitro screening of the drug substance in a test system that is specific for humans. The antibodies can be used in immunohistochemistry and for diagnostic and therapeutic applications. The present sequence represents a human neuronal NACHR alpha7 subunit encoding cDNA

Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Query Match 12.8%; Score 296.2; DB 6; Length 1876;
Best Local Similarity 60.7%; Pred. No. 3.3e-72;
Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;

QY 911 AAGTTTGCTAGCAGGATATCATGAAAGAGAGCTGTTACAGCATCTTTGGATCCTTATA 970
Db 125 ACGTGTCCCTGCAAGCGAGTTCAGAGGAGCTTTTACAGGAGCTGCTCAAGACTACA 184
QY 971 ATACACTAGAACCTCCCGTTCTCAATGAATCGGACCCGTTTACAAATTAAGCTTTGGTTAA 1030
Db 185 ATCCCTTGGAGAGGCCGTGGCCATGACTCGCAACCACTCACCGTCTACTTCTCCCTGA 244
QY 1031 CTTTAATGCAATATTCGATGTGGACGAGAAAAATCAATGTCTAGTCACTAATGTGTGT 1090
Db 245 GCCTCTCGAGATCATGGACGTGGATGAGAAACCAAGTTTAAACCAACATTTGGC 304
QY 1091 TAAACTGGAGTGGAAACGACATGAATCTCCGCTGGAAACACCTCCGACTATGGCGAGTTA 1150
Db 305 TSCAAATGCTTTGGACAGATCACTATTACAGTGGATGTGTCAGATATCCAGGGTGA 364
QY 1151 AGGATCTCGGAATACCGCGCATTCGCACTGGAAGCCGAGCTGCTGATGTACAACTGTG 1210
Db 365 AGACTGTTCGTTTCCAGATGGCCAGATTTGGAAACCAAGATTTCTCTATAACAGTG 424
QY 1211 CGGATCAGGGATTTGACCGCACCTACCAGCAAGTGTGTGGTGGGAAACACGCTCGT 1270

Db 425 CTGATGAGCGCTTTGAGCCACATTCACACTAACGTTGGTGAATTCCTCTGGGCATT 484
Qy 1271 GTCTATACGTTCCGCCGGGATCTTCAAGTCGAGTCGACATCGACATCACGTGGTTCC 1330
Db 485 GCCAGTACCTGCTCCAGGCATATTCAGAGTTCCTGTACATCATGATGAGCTGGTTTC 544
Qy 1331 CCTTCGATGACCGGCTGCGAGATGAAGTTGGCAGTTGGACCTACGACGATTCCAGC 1390
Db 545 CCTTGATGTCACACATGCAAACTGAAGTTGGTCTGTGCTCTACGGAGCTGGTCCT 604
Qy 1391 TGGATTTACAAATTACAAAGATGAAGTGGCGGTGATATCAGCAGTTACGTCTAACGGCG 1450
Db 605 TGGATCTGCAGATCAGGAGCA-----GATATCAGTGGCTATATCCCAATGGAG 655
Qy 1451 AGTGGAACTACTGGGTGTGCCGGGAAACGTAACGAGATCTATTACAACTGCTGCCCGG 1510
Db 656 AATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGGTTCTATGAGTGTGCAAAAG 715
Qy 1511 AACCTATATAGACATCACCTTCGCCATCATCCGCCGACGAACTGTACTATTCT 1570
Db 716 AGCCCTACCCGATGTACCTTCAAGTGACCATGGCCGCGAGGACGCTCTACTATGGCC 775
Qy 1571 TCAACCTGATCATACCTTGTGTTACTGATTCCTCCATGGCCCTTGTCTGGATTACCCCTGC 1630
Db 776 TCAACCTGCTGATCCCTGTGTCTCATCTCCGCCCTCGCCCTGTGGTGTCTCTGCTTC 835
Qy 1631 CGCCAGATTCCGGTGAAAAATTATCGTGGGTGTACCATCTTGTCTCTCGTGACCGTGT 1690
Db 836 CTGCAGATTCCGGGGAGAAGATTTCCCTGGGGATAACAGTCTTACTCTCTTACCGTCT 895
Qy 1691 TTCTGAATATGTTGCCGAGACAATGCCGGCTACTTCCGATGCGGTGCCATTG 1743
Db 896 TCAATGCTGCTGCTGAGATCATGCCGCAACATCCGATTCCGTACCATTG 948

Search completed: May 7, 2004, 15:01:30
Job time : 939.055 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:08:32 ; Search time 6207.93 Seconds
(without alignments)

11111.850 Million cell updates/sec

Title: US-09-303-232-1_COPY_372_2681

Perfect score: 2310

Sequence: 1 atgaataatgcacaaactgaa.....caccacatatattgtctcg 2310

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	424.8	18.4	885	12	BG632919	BG632919 GH16126.3
2	332.2	14.4	1101	29	CNS000IF	AL058211 Drosophil
3	331.2	14.3	1007	29	CNS00HJU	AL073676 Drosophil
4	323.8	14.0	978	29	CNS006F9	AL064281 Drosophil

5	281.8	12.2	1201	9	AL530299	AL530299
6	265	11.5	607	9	AI292581	AI292581 GH15518.5
7	251	10.9	2296	14	CD013901	CD013901 90134548
8	235.2	10.2	922	13	BU915857	BU915857 AGENCOURT
9	232.6	10.1	2940	11	AK034228	AK034228 Mus muscu
10	232.6	10.1	3230	11	AK083157	AK083157 Mus muscu
11	228.8	9.9	833	14	CB245337	CB245337 UI-M-FY0-
12	225.4	9.8	1436	29	AY402873	AY402873 Homo sapi
13	222.2	9.6	1436	29	AY402874	AY402874 Pan trogl
14	219.6	9.5	755	13	BU702422	BU702422 UI-M-FI0-
15	219.2	9.5	1374	29	AY406232	AY406232 Mus muscu
16	219.2	9.5	4290	11	AK029177	AK029177 Mus muscu
17	218.6	9.5	730	14	CF742344	CF742344 UI-M-HB0-
18	217	9.4	615	14	CB149460	CB149460 K-BST0205
19	215.8	9.3	1436	29	AY402875	AY402875 Mus muscu
20	215.8	9.3	1864	11	AK053497	AK053497 Mus muscu
21	215.8	9.3	2916	11	AK051730	AK051730 Mus muscu
22	215.8	9.3	3126	11	AK080415	AK080415 Mus muscu
23	207	9.0	3483	11	AK081254	AK081254 Mus muscu
24	207	9.0	4037	11	AK049722	AK049722 Mus muscu
25	207	9.0	4046	11	AK051742	AK051742 Mus muscu
26	206.4	8.9	683	12	BM639954	BM639954 170006876
27	205.4	8.9	908	13	BU149265	BU149265 AGENCOURT
28	203.4	8.8	902	29	AY407186	AY407186 Mus muscu
29	202.8	8.8	411	12	BI516733	BI516733 B160023A
30	202.2	8.8	2010	12	BI516843	BI516843 B160023B
31	201.8	8.7	410	12	BI516843	BI516843 B160023B
32	200	8.7	658	12	BM711715	BM711715 UI-E-CL1-
33	199.6	8.6	902	29	AY407184	AY407184 Homo sapi
34	198	8.6	755	10	AW914206	AW914206 EST345510
35	197.2	8.5	1036	13	EX437801	EX437801 BX437801
36	196.8	8.5	1374	29	AY406230	AY406230 Homo sapi
37	196.8	8.5	1466	14	CD013891	CD013891 90139069
38	196.6	8.5	2513	11	AK033068	AK033068 Mus muscu
39	194.4	8.4	1603	29	AY411327	AY411327 Homo sapi
40	194	8.4	759	14	CD804155	CD804155 UI-M-GV0-
41	191.4	8.3	791	14	CF739825	CF739825 UI-M-HD0-
42	189.6	8.2	720	14	CA373069	CA373069 647093 NC
43	188.2	8.1	1374	29	AY406231	AY406231 Pan trogl
44	186.4	8.1	677	14	CF919760	CF919760 Bf1or531.
45	183.4	7.9	1442	29	AY402876	AY402876 Homo sapi

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS
DEFINITION
GH16126.3prime GH Drosophila melanogaster head pOT2 Drosophila
melanogaster cDNA clone GH16126 3 similar to CG4128; Fban0004128
'ion channel', located on: 2L 30D1-30E1;: 04/10/2001, mRNA sequence.
ACCESSION
BG632919
VERSION
BG632919.1
KEYWORDS
EST.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 885)
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
Other ESTs: GH16126.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or

more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003511: arm:x [18792641,19136447]
estimated-cyto:1BA3-18C6: 04/10/2001
Plate: GH.161 row: C column: 2
High quality sequence stop: 784.

FEATURES

Location/Qualifiers
1..885
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH16126"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/clone_lib="GH Drosophila melanogaster head pOT2"
/note="Organ: head; Vector: pOT2; Site: 1; EcorI; Site 2;
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."

ORIGIN

Query Match 18.4%; Score 424.8; DB 12; Length 885;
Best Local Similarity 74.6%; Pred. No. 1.8e-92;
Matches 534; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1028 TAACCTTAATGCAATATTCAGATGTCGACGAGAAATCAATTCGTAGTCACTAATGTGT 1087
DB |||||
DB 884 TAACACTCATGACATTCATGTCGACGAGAAATCAATTCGTATTAACGAATATT 825
QY 1088 GGTAAACCTGAGTGGAAACACATGAATCTCCGCTGGAAACACCTCCGACTATGCGGAG 1147
DB |||||
DB 824 GGCTCAATTCGATGGAACGATATGATCTTCGATGAAATTCGATGAGTTCGGTGTG 765
QY 1148 TTAAGGATCTCGAATACCGCGCATCGCATCTGGAACCGGACGTCGTGATGTACAACA 1207
DB |||||
DB 764 TCGCGATCTCGAATTCGCGCACATCGCTATGGAACCGGATCTACTGATGTACAACA 705
QY 1208 GTCCGATGAGGATTTGACCGCACTACAGACGACGTCGTGTGCGGAACACGCT 1267
DB |||||
DB 704 GTGCGGACGAGGGCTTCATGGAACGATGACGCCAANAATGTGTGTGTCGCAATATGGGA 645
QY 1268 CGTGTCTATAGCTTCGCGCGGGGATCTTCAAGTCGACGTCGCAAGATCGACATCAGTGT 1327
DB |||||
DB 644 GCTGTCTGATGACGCGCAGGTATATTAAAGTCAACGTTGTAAGATCGACATTAAGTGT 585
QY 1328 TCCCTTCGATGACAGCGGTGCGAGATGAAGTTCGGCAGTGTGACCTACGACGATTC 1387
DB |||||
DB 584 TTCCATTCGACGATCAGAGATGTGAATGAAATTTGGTTCGTGGACCTACGATGGGTTTC 525
QY 1388 AGCTGATTTACAATTACAAGATGAATGCGCGGTGATCAGCAGTTACGTCCTCAACG 1447
DB |||||
DB 524 AGTTGACCTGAGTTGAGGACGAGAGCTGGTGGCGACATTTCTAGCTTTATACCAATG 465
QY 1448 GCGAGTGGAACTACTGGGTGTGCCCGCAACGTPAACGAGATCTTATTACAACTCTGTC 1507
DB |||||
DB 464 GCGAATGGACTTGTAGTGTGCCCGGTAAACGAAATGAAATCTACTATAATTGCTGCC 405
QY 1508 CGGAACCTTATAGACATCACCCTTCGCCATCATCATCGCCGACGAACACATGTAATT 1567
DB |||||
DB 404 CAGAACCTTATATTGACATAACATTCGCCATTTTGAATAGGCGCAAAAGCTTGTACTATT 345
QY 1568 TCTTCAACCTGATCATCTGTGATGATTCCTTCCATGCGCTTGTCTCGGATTTCAACC 1627
DB |||||
DB 344 TTTTCAATCTGATGTGCGGTGCGTACTGATCGCTTCCATGCGACTGCTAGGTTTACAC 285
QY 1628 TGCCTCCAGATTCGGGTGAAAAATTTATCGTGGGTGTTTACCAATCTTGTCTCTCGTGACCG 1687
DB |||||
DB 284 TGCACCAAGATTCGTGTGAAAGCTTTCGTTGGAGTTACAATTTCTATTATCGCTTACAG 225
QY 1688 TGTTCCTGAATGTTGCGGAGCAATGCCGGCTACTTCCGATCGGTGCGCATTTG 1743
DB |||||
DB 224 TCTTCTCAACATGTTGGCGGCAACAAATGCCGGCGACCTCCGATCGGTACCGCTG 169

RESULT 2

LOCUS

DEFINITION

CNS00001F 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR02C08 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

TITLES

AUTHORS

JOURNAL

COMMENT

AL058211 GI:4930650

GSS.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osogawa and

Aaron Mammoser in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

ECORI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain Y2; cn bw sp, the same strain used for the BDGP's

p1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR02C08"

/clone_lib="RPCI-98"

/note="end : T7"

ORIGIN

Query Match 14.4%; Score 332.2; DB 29; Length 1101;
Best Local Similarity 98.5%; Pred. No. 8.2e-70; Indels 0; Gaps 0;
Matches 334; Conservative 1; Mismatches 4;

QY 1053 GGACGAGAAAAATCAATTCGTAGTCACCTAATGTGTGTTAAACCTGGAGTGAACGACAT 1112
DB |||||
DB 113 GGACGAGAAAAATCAATTCGTAGTCACCTAATGTGTGTTAAACCTGGAGTGAACGACAT 172
QY 1113 GAATCTCCGTGGAAACACCTCCGACTATGGCGAGTTAAGGATCTGCGAATACCCCGCA 1172
DB |||||
DB 173 GAATCTCCGTGGAAACACCTCCGACTATGGCGAGTTAAGGATCTGCGAATACCCCGCA 232
QY 1173 TCGCATCTGGAGCGGACGTCGTGATGTACAACTGCGGATGAGGATTTGACGGCAC 1232
DB |||||
DB 233 TCGCATCTGGAGCGGACGTCGTGATGTACAACTGCGGATGAGGATTTGACGGCAC 292
QY 1233 CTACCAAGCGAATCGAAGATCGACATCAGTGTTCCTTCCTTCGATGATGACCGGAT 1292
DB |||||
DB 293 CTACCAAGCGAATCGAAGATCGAATCAACGCTCGTGTCTATACGTTCCGCGGGAT 352
QY 1293 CTTCAAGTCGAGCTGCAAGATCGACATCAGTGTTCCTTCCTTCGATGATGACCGGATGCGCA 1352
DB |||||
DB 353 CTTCAAGTCGAGCTGCAAGATCGACATCAGTGTTCCTTCCTTCGATGATGACCGGATGCGCA 412
QY 1353 GATGAAGTTCGGCAGTTGACCTACGAGGATTTCCAGCT 1391
DB |||||
DB 413 GATGAAGTTCGGCAGTTGACCTACGAGGATTTCCAGCT 451

```

Db
425 GATGAAGTTCGGCAGTTGGACCTAGACGAGTTCAGGT 463

RESULT 4
CNS000HJTU
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR35F05 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL073676
AL073676.1 GI:4953355
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1007)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .1007
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR35F05"
/clone_lib="RPCI-98"
/note="end : TET3"

FEATURES
source
ORIGIN
Query Match 14.3%; Score 331.2; DB 29; Length 1007;
Best Local Similarity 98.2%; Pred. No. 1.4e-69;
Matches 333; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1053 GGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAAACGACAT 1112
DB 125 GGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAAACGACAT 184
QY 1113 GAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGGAATACCGCGCA 1172
DB 185 GAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGGAATACCGCGCA 244
QY 1173 TGCATCTCGAGCCGGAGCTGTGTGATGTACAACAGTGGGATGAGGGATTTCACGGCAC 1232
DB 245 TGCATCTCGAGCCGGAGCTGTGTGATGTACAACAGTGGGATGAGGGATTTCATGGCAC 304
QY 1233 CTACACAGACGAGCTGGTGGGACCAACCGCTCGTGTCTATACGTTCCGCGGGGAT 1292
DB 305 CTACACAGACGAGCTGGTGGGACCAACCGCTCGTGTCTATACGTTCCGCGGGGAT 364
QY 1293 CTTCAAGTCGACGTGCAAGATCGACATCACGTGTTCCCTTCGATGACACCGGTGGCA 1352
DB 365 CTTCAAGTCGACGTGCAAGATCGACATCACGTGTTCCCTTCGATGACACCGGTGGCA 424
QY 1353 GATGAAGTTCGGCAGTTGGACCTAGACGAGTTCAGGT 1391

```

```

Db
425 GATGAAGTTCGGCAGTTGGACCTAGACGAGTTCAGGT 463

RESULT 4
CNS0006F9
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR13K21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL064281
AL064281.1 GI:4944356
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 978)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .978
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR13K21"
/clone_lib="RPCI-98"
/note="end : TET3"

FEATURES
source
ORIGIN
Query Match 14.0%; Score 323.8; DB 29; Length 978;
Best Local Similarity 99.1%; Pred. No. 8.9e-68;
Matches 336; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1053 GGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAAACGACAT 1112
DB 117 GGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAAACG-CAT 175
QY 1113 GAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGGAATACCGCGCA 1172
DB 176 GAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGGAATACCGCGCA 235
QY 1173 TGCATCTCGAGCCGGAGCTGTGTGATGTACAACAGTGGGATGAGGGATTTCACGGCAC 1232
DB 236 TGCATCTCGAGCCGGAGCTGTGTGATGTACAACAGTGGGATGAGGGATTTCATGGCAC 295
QY 1233 CTACACAGACGAGCTGGTGGGACCAACCGCTCGTGTCTATACGTTCCGCGGGGAT 1292
DB 296 CTACACAGACGAGCTGGTGGGACCAACCGCTCGTGTCTATACGTTCCGCGGGGAT 355
QY 1293 CTTCAAGTCGACGTGCAAGATCGACATCACGTGTTCCCTTCGATGACACCGGTGGCA 1352
DB 356 CTTCAAGTCGACGTGCAAGATCGACATCACGTGTTCCCTTCGATGACACCGGTGGCA 415
QY 1353 GATGAAGTTCGGCAGTTGGACCTAGACGAGTTCAGGT 1391

```

Db	416	GATGAAGTTCGGCAGTTGGACCTACGACGATCCAGGT	454
Db	470	ACAGTGTGATGAGCGCTTTGACGCCACATTCACACATTAACGTTGGTGAATCTTCTGT	529
Qy	1265	GCTCGTGTCTATACGTTCCCGCGGGATCTTCAAGTCGAGTCGACGATCGACATACAGT	1324
Db	530	GGCATTCGCCAGTWCCTGCCTCCAGGCATATTCAGAGTTCTGCTACATCGATACGCT	589
Qy	1325	GGTTCCCTTCGATACCAAGCGGTGCGAGATGAAGTTCCGGAGTTGGACCTACGACGAT	1384
Db	590	GGTTCCCTTTGATGTGCAGCACTGCAAACTGAAGTTTGGTCTCTTACGGAGGCT	649
Qy	1385	TCAGCTGGATTACATATTACAGATGAACCTGCGGTGATATCAGCAGTTTACCTGTCTCA	1444
Db	650	GGTCTTTGGATGTCAGATGAGAGGCA	700
Qy	1445	ACGCGAGTGGGAACCTACTGGGTGTGCCGCGCAACGTAACGAGATCTATTACAACCTGT	1504
Db	701	ATGAGAGATGGGACCTAGTGGGAATCCCGGCAAGAGAGTGAAGGTTCCTATGAGTCT	760
Qy	1505	GCCCGGAACCTATATAGACATCACTTGGCATCATCCGCGGAGGAGCACTGTACT	1564
Db	761	GCAAGAGCCCTACCCGATGTCACTTCAAGTGACCATGCGCGCGAGGACGCTCTACT	820
Qy	1565	ATTCTTTCAACCTGATCATACCTTTGTCTACTGATTGCTCCATGCGCTTGTCTCGATTCA	1624
Db	821	ATGGCTCAACCTGCTGATCCCTGTGTCTCACTCCGCTCGCTCGCTGTGTGTCTCT	880
Qy	1625	CCCTGCCGCGCAGATTC--GGGTGAAAATTTATCGCTGGTGTATACCATTTTGTCTCGCT	1682
Db	881	TGCTTCTCGAGATTCGGGGAGAGATTCCCTKGGGATAACAGTATTACTCTCTCT	940
Qy	1683	GACCGTGTCTTGAATATGGTTGCGGAGACAATGCGGCTACTTCCGATGGGTGCCATT	1742
Db	941	TACCGTCTTCATGCTGCTGCTGCTGATCATGCCGCAACATCCGATTCGGTACCAT	1000
Qy	1743	G 1743	
Db	1001	G 1001	
RESULT 6			
LOCUS	AI292581	607 bp	linear
DEFINITION	GH15518.5prime GH Drosophila melanogaster head P012 Drosophila melanogaster cDNA clone GH15518 5 similar to C04128: FBan004128 'ion channel', located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.		
ACCESSION	AI292581		
VERSION	AI292581.1	GI:3941988	
KEYWORDS	EST.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
REFERENCE	Harvey,D., Broksstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.		
AUTHORS	BDGP/HHMI Drosophila EST Project		
TITLE	Unpublished (2001)		
JOURNAL	Other_ESTS: GH15518.3prime		
COMMENT	Contact: Stapleton, M.		
BDGP			
Lawrence Berkeley National Lab			
One Cyclotron Rd, Berkeley, CA 94720, USA			
Fax: 510 486 6798			
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu			
hit genomic AE003626; arm:2L [9617316,9882551]			
estimated-cyto:30C7-30F4: 04/10/2001			
Plate: GH.155 row: B column: 6			
High quality sequence stop: 521			
POLYA=No.			
FEATURES			
Location/Qualifiers			
1..607			
/organism="Drosophila melanogaster"			
source			
Query Match	12.2%;	Score 281.8;	DB 9;
Best Local Similarity	59.8%;	Pred. No. 1.7e-57;	Length 1201;
Matches 503;	Conservative	5;	Mismatches 322;
		Indels	11;
		Gaps	2;
Qy	905	CTGCTAAAGTTGGCTAGCAGATATCATGAAAAGAGACTGTACAGATCTTTGGATC	964
Db	170	CTCTCGCAGTCTCCCTGCAAGCGAGTWCACAGGAAGCTTTACAAAGGAGCTGGTCAAGA	229
Qy	965	CTTATAATACACTAGACGTCCTGCTCTCAATGAATCGGACCGGTACAAATTAAGCTTG	1024
Db	230	ACTAAATCCCTTGAGAGCGCCGCGGCAATGACTGCAACACTCACCGTCTACTTCT	289
Qy	1025	GTTTAACCTTAAATGCAATTTATCGATGTGGACGAGAAAATCAATTCGTAGTCACTAATG	1084
Db	290	CCCTGAGCTCTCCGATCATGGACGTGGATGAGAGAACAAGTTTAAACCAACA	349
Qy	1085	TGTGTTAAACTGGAGTGAAGACGATCAATTCGCTGGGACACCTCCGACTATGGCG	1144
Db	350	TTTGGCTGCAAAATGCTTGGACAGATCACTATTACAGTGAATGTGTGAGATATCCAG	409
Qy	1145	GAGTTAAGGATCTGCGAATACCGCCATCGCATCTGGAAGCGCGGACGTGTGATGACA	1204
Db	410	GGGTGAAGACTGTTCGTTTCCAGATGGCCAGATTGGAAACAGACATTTCTCTATA	469
Qy	1205	ACAGTCGGATGAGGATTTGACGGCACTTACAGACGAACGTGGTGGTGGCAACAG	1264

```
/mol_type="mrna"
/db_xref="taxon:7227"
/clone="GH15518"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/clone_lib="GH Drosophila melanogaster head pot2"
/note="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2:
XhoI; Sized fractionated cDNAs were directly ligated into
pot2. Plasmid cDNA library."
```

ORIGIN

```
Query Match 11.5%; Score 265; DB 9; Length 607;
Best Local Similarity 71.4%; Pred. No. 1.7e-53;
Matches 349; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 909 TAAAGTTTGGCTACAGATATCATGAAAGAGACTGTTACAGATCTTTTGGATCCTTA 968
D 909 TAAAGTTTGGCTACAGATATCATGAAAGAGACTGTTACAGATCTTTTGGATCCTTA 968
D 909 TAAAGTTTGGCTACAGATATCATGAAAGAGACTGTTACAGATCTTTTGGATCCTTA 968
QY 969 TAATACACTAGAACGTCCTTCTCAATGAATCGGACCCGTTTACAAATTAAGCTTTGGTT 1028
D 969 TAATACACTAGAACGTCCTTCTCAATGAATCGGACCCGTTTACAAATTAAGCTTTGGTT 1028
D 969 TAATACACTAGAACGTCCTTCTCAATGAATCGGACCCGTTTACAAATTAAGCTTTGGTT 1028
QY 1029 AACTTTAATGCAATTATTCGATGTGGACGAGAAATCAATTGCTAGTCACTAATGTGTG 1088
D 1029 AACTTTAATGCAATTATTCGATGTGGACGAGAAATCAATTGCTAGTCACTAATGTGTG 1088
D 1029 AACTTTAATGCAATTATTCGATGTGGACGAGAAATCAATTGCTAGTCACTAATGTGTG 1088
QY 1089 GTTAAACTGAGTGGAGACGATGATCTCCGCTGGAAACACCTCCGACTATGGCGAGT 1148
D 1089 GTTAAACTGAGTGGAGACGATGATCTCCGCTGGAAACACCTCCGACTATGGCGAGT 1148
D 1089 GTTAAACTGAGTGGAGACGATGATCTCCGCTGGAAACACCTCCGACTATGGCGAGT 1148
QY 286 GCTTTTCGTTGAGTGGAAACGACTACAAATCTCGCTGGAATGAAACGGAATACGCGGGT 345
D 286 GCTTTTCGTTGAGTGGAAACGACTACAAATCTCGCTGGAATGAAACGGAATACGCGGGT 345
D 286 GCTTTTCGTTGAGTGGAAACGACTACAAATCTCGCTGGAATGAAACGGAATACGCGGGT 345
QY 1149 TAAGGATCTCGAATACCGCGCATCGCATCTGGAACGCGAGCTGCTGATGTGTACACAG 1208
D 1149 TAAGGATCTCGAATACCGCGCATCGCATCTGGAACGCGAGCTGCTGATGTGTACACAG 1208
D 1149 TAAGGATCTCGAATACCGCGCATCGCATCTGGAACGCGAGCTGCTGATGTGTACACAG 1208
QY 346 CAGGATCTAGCAATACAGCCCAACAGCTGTGAAGCCGAGCTGCTCATGTACAAACAG 405
D 346 CAGGATCTAGCAATACAGCCCAACAGCTGTGAAGCCGAGCTGCTCATGTACAAACAG 405
D 346 CAGGATCTAGCAATACAGCCCAACAGCTGTGAAGCCGAGCTGCTCATGTACAAACAG 405
QY 1209 TCGGATAGGGATTTGACGGCACCTACAGACGAGCTGTGTGTGGGAAACACGGCTC 1268
D 1209 TCGGATAGGGATTTGACGGCACCTACAGACGAGCTGTGTGTGGGAAACACGGCTC 1268
D 1209 TCGGATAGGGATTTGACGGCACCTACAGACGAGCTGTGTGTGGGAAACACGGCTC 1268
QY 406 CGCGGATAGGGATTTGACGGCACCTACAGACGAGCTGTGTGTGGGAAACACGGCTC 1328
D 406 CGCGGATAGGGATTTGACGGCACCTACAGACGAGCTGTGTGTGGGAAACACGGCTC 1328
D 406 CGCGGATAGGGATTTGACGGCACCTACAGACGAGCTGTGTGTGGGAAACACGGCTC 1328
QY 1269 GTCTCTATAGTTCGCGCGGGATCTTCAAGTGCAGCTGCAAGTGCAGTCACTACAGTGGTT 1328
D 1269 GTCTCTATAGTTCGCGCGGGATCTTCAAGTGCAGCTGCAAGTGCAGTCACTACAGTGGTT 1328
D 1269 GTCTCTATAGTTCGCGCGGGATCTTCAAGTGCAGCTGCAAGTGCAGTCACTACAGTGGTT 1328
QY 466 TTGTCTGTAGTCCGCTGCTATCTTCAAGAGCACATGCAAGTGCAGTCACTACAGTGGTT 525
D 466 TTGTCTGTAGTCCGCTGCTATCTTCAAGAGCACATGCAAGTGCAGTCACTACAGTGGTT 525
D 466 TTGTCTGTAGTCCGCTGCTATCTTCAAGAGCACATGCAAGTGCAGTCACTACAGTGGTT 525
QY 1329 CCCCTTCGATACAGCGGTGCGAGATGAAGTTCGCGAGTGGACCTACAGCGAGTCCA 1388
D 1329 CCCCTTCGATACAGCGGTGCGAGATGAAGTTCGCGAGTGGACCTACAGCGAGTCCA 1388
D 1329 CCCCTTCGATACAGCGGTGCGAGATGAAGTTCGCGAGTGGACCTACAGCGAGTCCA 1388
QY 526 CCATTGTGATGACCAATGCGGAATGAAATTCGGTAGTGGACCTTACGATGGAATCA 585
D 526 CCATTGTGATGACCAATGCGGAATGAAATTCGGTAGTGGACCTTACGATGGAATCA 585
D 526 CCATTGTGATGACCAATGCGGAATGAAATTCGGTAGTGGACCTTACGATGGAATCA 585
QY 1389 GCTGGATTT 1397
D 1389 GCTGGATTT 1397
D 1389 GCTGGATTT 1397
D 586 GTTGGATTT 594
```

RESULT 7

```
CD013901
LOCUS 2296 bp mrna linear EST 21-OCT-2003
DEFINITION Single gene library Homo sapiens cDNA, mRNA sequence.
ACCESSION CD013901
VERSION CD013901.1 GI:37777431
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2296)
Jin, P., Fu, G.K., Wilson, A.D., Yang, J., Chien, D., Hawkins, P.R.,
Au-Yang, J. and Stuve, L.I.
PCR isolation and cloning of novel splice variant mRNAs from known
drug target genes
Unpublished (2003)
Contact: Jin, P.
Incyte Corporation
3160 Porter Drive, Palo Alto, CA 94304, USA
```

```
Tel: 650 621 8639
Fax: 650 621 8965
Email: pjin@incyte.com.
```

FEATURES

source

Location/Qualifiers

```
1..2296
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone_lib="Single gene library"
/note="Vector: pDrive Cloning Vector; RT-PCR was performed
using gene-specific primers flanking the open-reading
frame. PCR products were subcloned into pDrive Cloning
Vector and sequenced completely using M13 forward and
reverse primers. Sequencing gaps were closed by
re-sequencing using primers flanking the gapped areas."
```

ORIGIN

```
Query Match 10.9%; Score 251; DB 14; Length 2296;
Best Local Similarity 58.1%; Pred. No. 7.2e-50;
Matches 462; Conservative 0; Mismatches 330; Indels 3; Gaps 1;

QY 933 TGAAGAGAGACTGTTACAGATCTTTTGGATCCTTATAATACACTAGAACGTCGCTTCT 992
D 933 TGAAGAGAGACTGTTACAGATCTTTTGGATCCTTATAATACACTAGAACGTCGCTTCT 992
D 933 TGAAGAGAGACTGTTACAGATCTTTTGGATCCTTATAATACACTAGAACGTCGCTTCT 992
QY 993 CAATGAATCGGACCCGTTACAAATTAAGCTTTGGTTTAACTTTAATGCAATTAATCGATG 1052
D 993 CAATGAATCGGACCCGTTACAAATTAAGCTTTGGTTTAACTTTAATGCAATTAATCGATG 1052
D 993 CAATGAATCGGACCCGTTACAAATTAAGCTTTGGTTTAACTTTAATGCAATTAATCGATG 1052
QY 409 CAACACTTCAGACGTTGGTGAATGTGCGCTTTGGACTGTCCATCGCTCAGCTCATCGATG 468
D 409 CAACACTTCAGACGTTGGTGAATGTGCGCTTTGGACTGTCCATCGCTCAGCTCATCGATG 468
D 409 CAACACTTCAGACGTTGGTGAATGTGCGCTTTGGACTGTCCATCGCTCAGCTCATCGATG 468
QY 1053 GAGCAGAGAAAACCAATGCTAGTCACTAATGTGTGTGTTAAACTTGAAGTGAACGACAT 1112
D 1053 GAGCAGAGAAAACCAATGCTAGTCACTAATGTGTGTGTTAAACTTGAAGTGAACGACAT 1112
D 1053 GAGCAGAGAAAACCAATGCTAGTCACTAATGTGTGTGTTAAACTTGAAGTGAACGACAT 1112
QY 469 GGATGAGAGAACCAATGATGACCAACAGCTTGGCTAAACAGAGTGGAGCGACTA 528
D 469 GGATGAGAGAACCAATGATGACCAACAGCTTGGCTAAACAGAGTGGAGCGACTA 528
D 469 GGATGAGAGAACCAATGATGACCAACAGCTTGGCTAAACAGAGTGGAGCGACTA 528
QY 1113 GAATCTCCGCTGGAAACACCTCCGACTATGGCGAGTTAAGGATCTGCGAATACCGCGCA 1172
D 1113 GAATCTCCGCTGGAAACACCTCCGACTATGGCGAGTTAAGGATCTGCGAATACCGCGCA 1172
D 1113 GAATCTCCGCTGGAAACACCTCCGACTATGGCGAGTTAAGGATCTGCGAATACCGCGCA 1172
QY 529 CAAACTGCGCTGGAAACCCCGCTGATTTTGGCAACATCACATCTCTAGGGTCCCTCTGA 588
D 529 CAAACTGCGCTGGAAACCCCGCTGATTTTGGCAACATCACATCTCTAGGGTCCCTCTGA 588
D 529 CAAACTGCGCTGGAAACCCCGCTGATTTTGGCAACATCACATCTCTAGGGTCCCTCTGA 588
QY 1173 TCGCATCTGGAACCGGACGCTGCTGATGATACACAGTGGCGGATGAGGATTTGACGGCAC 1232
D 1173 TCGCATCTGGAACCGGACGCTGCTGATGATACACAGTGGCGGATGAGGATTTGACGGCAC 1232
D 1173 TCGCATCTGGAACCGGACGCTGCTGATGATACACAGTGGCGGATGAGGATTTGACGGCAC 1232
QY 589 GATGATCTGGATCCCGACATTTCTTCTACAAACATGACAGTGGGAGTTTGCAGTGAC 648
D 589 GATGATCTGGATCCCGACATTTCTTCTACAAACATGACAGTGGGAGTTTGCAGTGAC 648
D 589 GATGATCTGGATCCCGACATTTCTTCTACAAACATGACAGTGGGAGTTTGCAGTGAC 648
QY 1233 CTACACAGAACGTTGGTGGCGGACACGCTGCTGTCTATACGTTCCGCGGGGAT 1292
D 1233 CTACACAGAACGTTGGTGGCGGACACGCTGCTGTCTATACGTTCCGCGGGGAT 1292
D 1233 CTACACAGAACGTTGGTGGCGGACACGCTGCTGTCTATACGTTCCGCGGGGAT 1292
QY 649 CCACATGACCAAGCGCCACCTCTTCTCCACGGCAGCTGTGCTGCTGCGCGGCGCAT 708
D 649 CCACATGACCAAGCGCCACCTCTTCTCCACGGCAGCTGTGCTGCTGCTGCGCGGCGCAT 708
D 649 CCACATGACCAAGCGCCACCTCTTCTCCACGGCAGCTGTGCTGCTGCTGCGCGGCGCAT 708
QY 1293 CTTCAAGTGCAGTGCAGATCGACATCGACATCGCTGCTTCCCCTTCGATGACCGCGTCCGA 1352
D 1293 CTTCAAGTGCAGTGCAGATCGACATCGACATCGCTGCTTCCCCTTCGATGACCGCGTCCGA 1352
D 1293 CTTCAAGTGCAGTGCAGATCGACATCGACATCGCTGCTTCCCCTTCGATGACCGCGTCCGA 1352
QY 709 CTACAAAGAGCTCTGCAGCATCGACGTCACCTTCTTCCCTTCGACCCAGCACTGCAA 768
D 709 CTACAAAGAGCTCTGCAGCATCGACGTCACCTTCTTCCCTTCGACCCAGCACTGCAA 768
D 709 CTACAAAGAGCTCTGCAGCATCGACGTCACCTTCTTCCCTTCGACCCAGCACTGCAA 768
QY 1353 GATGAAGTTCCGCGAGTTGGACCTACGACGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 1412
D 1353 GATGAAGTTCCGCGAGTTGGACCTACGACGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 1412
D 1353 GATGAAGTTCCGCGAGTTGGACCTACGACGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 1412
QY 769 GATGAAGTTCCGCGAGTTGGACCTACGACGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 828
D 769 GATGAAGTTCCGCGAGTTGGACCTACGACGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 828
D 769 GATGAAGTTCCGCGAGTTGGACCTACGACGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 828
QY 1413 AACTGGCGGTGATATCAGCAGTTTACGCTTCAACGCGAGTGGGAACTACTGGGTGGCC 1472
D 1413 AACTGGCGGTGATATCAGCAGTTTACGCTTCAACGCGAGTGGGAACTACTGGGTGGCC 1472
D 1413 AACTGGCGGTGATATCAGCAGTTTACGCTTCAACGCGAGTGGGAACTACTGGGTGGCC 1472
QY 829 GACTGT---GGACCTGAGGACTACTGGGAGAGCGGCGAGTGGGCCATCGTCAATGCCAC 885
D 829 GACTGT---GGACCTGAGGACTACTGGGAGAGCGGCGAGTGGGCCATCGTCAATGCCAC 885
D 829 GACTGT---GGACCTGAGGACTACTGGGAGAGCGGCGAGTGGGCCATCGTCAATGCCAC 885
QY 1473 CGGCAAACTGAACGAGATCTATTACAACTGCTGCCCGGAAACCCCTATATAGACATCACTT 1532
D 1473 CGGCAAACTGAACGAGATCTATTACAACTGCTGCCCGGAAACCCCTATATAGACATCACTT 1532
D 1473 CGGCAAACTGAACGAGATCTATTACAACTGCTGCCCGGAAACCCCTATATAGACATCACTT 1532
QY 886 GGGCACCCTTACAACAGCAAGAGTACGACTGTCTGCTGCGCGAGATCTACCCGACGTCACCTA 945
D 886 GGGCACCCTTACAACAGCAAGAGTACGACTGTCTGCTGCGCGAGATCTACCCGACGTCACCTA 945
D 886 GGGCACCCTTACAACAGCAAGAGTACGACTGTCTGCTGCGCGAGATCTACCCGACGTCACCTA 945
QY 1533 CGGCATCATCATCCGCGGACGAACTGATCTATTCTTCAACCTGATCATACCTTTGTGT 1592
D 1533 CGGCATCATCATCCGCGGACGAACTGATCTATTCTTCAACCTGATCATACCTTTGTGT 1592
D 1533 CGGCATCATCATCCGCGGACGAACTGATCTATTCTTCAACCTGATCATACCTTTGTGT 1592
QY 946 CGCCTTTCGTCATCCGCGGCTGCGCTCTTCTACACCATCAACCTCATCATCCCTGCT 1005
D 946 CGCCTTTCGTCATCCGCGGCTGCGCTCTTCTACACCATCAACCTCATCATCCCTGCT 1005
D 946 CGCCTTTCGTCATCCGCGGCTGCGCTCTTCTACACCATCAACCTCATCATCCCTGCT 1005
QY 1593 ACTGATTTGCTCCATGGCCCTTCTCGGATTCGATTCACCTGCGCGCAGATTCGGTGAATAAT 1652
D 1593 ACTGATTTGCTCCATGGCCCTTCTCGGATTCGATTCACCTGCGCGCAGATTCGGTGAATAAT 1652
D 1593 ACTGATTTGCTCCATGGCCCTTCTCGGATTCGATTCACCTGCGCGCAGATTCGGTGAATAAT 1652
QY 1006 GTCATCTCTGCTCAGTGTGCTGCTTCTTACCTGCGCTCCGACTCCGCTGCGCGAGAGAT 1065
D 1006 GTCATCTCTGCTCAGTGTGCTGCTTCTTACCTGCGCTCCGACTCCGCTGCGCGAGAGAT 1065
D 1006 GTCATCTCTGCTCAGTGTGCTGCTTCTTACCTGCGCTCCGACTCCGCTGCGCGAGAGAT 1065
QY 1653 ATCGCTGGGTGTTACCATCTTCTGCTGCTGACCGGTGTTCTGGAATATGTTGCCGAGAC 1712
D 1653 ATCGCTGGGTGTTACCATCTTCTGCTGCTGACCGGTGTTCTGGAATATGTTGCCGAGAC 1712
D 1653 ATCGCTGGGTGTTACCATCTTCTGCTGCTGACCGGTGTTCTGGAATATGTTGCCGAGAC 1712
QY 1066 CACGCTGTGCTATTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1125
D 1066 CACGCTGTGCTATTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1125
D 1066 CACGCTGTGCTATTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1125
```

```
QY 1713 AATGCCGGCTACTTC 1727
Db 1126 CATCCCGTCCACCTC 1140

RESULT 8
BU915857 922 bp mRNA linear EST 17-OCT-2002
LOCUS AGENCOURT 10492745 NICHID_XGC_001 Xenopus laevis cDNA clone
DEFINITION IMAGE:6642638 5', mRNA sequence.
ACCESSION BU915857
VERSION BU915857.1 GI:24097771
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
DEFINITION Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 922)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: ccapba-remail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Cloning Distribution: Agencourt Bioscience Corporation
Cloned through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUM14228 row: f column: 14
High quality sequence stop: 746.
Location/Qualifiers
1. 922
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6642638"
/tissue_type="oocytes"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHID_XGC_001"
/note="Vector: pCMV-SF0RT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.2 kb. Constructed by Life Technologies."
ORIGIN
Query Match 10.2%; Score 235.2; DB 13; Length 922;
Best Local Similarity 60.0%; Pred. No. 3.8e-46;
Matches 415; Conservative 0; Mismatches 268; Indels 9; Gaps 1;
QY 1051 GTGACGAGAAATCAATTCGTAGTCTAATGTGTGGTTAAACTGAGTGGACAC 1110
Db 37 GAGGATGAAAAAATCAAGTACTAACAACAACATATGCTGCTACAAATGCAATGGATAGAT 96
QY 1111 ATGATCTCCGTGGAACACCTCCGACTATGCGGAGTTAAGATCTGGATACCGCG 1170
Db 97 TACTATTTGCACTGGAACATGTTGAAATACCTCGGGGTAAATAATGTTCTGTTTCAAGT 156
QY 1171 CATCCCATCTGGAACCGGACGCTGTGTATGTATCAACAGTCGGATGAGGATTTGACGGC 1230
Db 157 GGACAGATTTGGAACCTGATATTCTTCTATACAGTCGAGATGATAGATTTGATGCC 216
QY 1231 ACCTACCGACGAACTGTGTGTGGGAAACACGCGCTGTGTCTATACGTTCCGCGGGG 1290
Db 217 TCCTTTTCACTAATGTTACTGTGTGAATCTCCAGTGACACTGCGCAATATCTGCCACGAGG 276
QY 1291 ATCTTCAAGTCGACGTGCAAGATCGACATCATCGTTGTTCCCTTCGATGACCGCGGTGC 1350
Db 277 ATTTTCAAAGTTCGTGTACATAGATGTCGATGTTCCCTTTTGACATACAGAAATG 336
QY 1351 GAGATGAAGTTCGGCAGTTGGACCTACGACGAGTCCAGCTGGATTTCAATTACAAGAT 1410

337 AACCTGAAGTTTGGCTCTTGACCTATGTGTGGATGGTCACTTGACCTGCAGATCTAGAA 396
1411 GAAACTGGCGGTGATATCAGCAGTTACGTCTCAACGCGAGTGGGAACCTACTCGGTGTG 1470
397 TCA-----GACATATCTGAATATATCGTAATGGAGAGTGGGACCTTGTAGCGGTT 447
1471 CCGGGAAAACGTACAGAGATCTATTACAACCTGCTGCCCGGAACCTATATAGACATCACC 1530
448 CTTGAAAAAGAGTGAAGATTACTATGAATGCTGCAAGGAACCTGACCTTGTATGTGACA 507
1531 TTGCCATCATCATCCCGCAGCAACTGTATCTATTCTTCAACCTGATCATACCTTGT 1590
508 TTTACGATAATAATGAGCGCAGGACTCTATACGACTTAACTTAAATTCCTTGT 567
1591 GTACTGATGCTCCATCGGCTGCTCGGATTCACCTGCCCGCAGATTCGGGTGAAAAA 1650
568 GTACTGATATCTGCTCTAGCCCTCTCTGGGTTCCTGCTTCAGCAGACTCAGGAGAGAG 627
1651 TTATCGCTGGGTGTACCATCTTCTCTCGCTGACCGGTTCCTGAATATGCTTGCAGAG 1710
628 ATCTCACTTGTATAACAGTCTTATTTCCCTCACTGTGTTTCATGCTTGTGTAGCTGAG 687
1711 ACAATGCCGCTACTTCCGATGCGGTGCCATT 1742
688 ATTATGCCAGCAACCTCAGACTCTGTGCCACT 719

RESULT 9
AK034228 2940 bp mRNA linear HTC 18-SEP-2003
LOCUS Mus musculus adult male dienecephalon cDNA, RIKEN full-length
DEFINITION enriched library, clone:930165116 product:cholinergic receptor,
nicotinic, alpha polypeptide 4, full insert sequence.
ACCESSION AK034228
VERSION AK034228.1 GI:26329798
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
DEFINITION Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wakahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
```

REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
AUTHORS		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
TITLE		Nature 420, 563-573 (2002)
JOURNAL		6 (bases 1 to 2940)
REFERENCE		Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashita, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saio, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
AUTHORS		Direct Submission
TITLE		Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
JOURNAL		cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
COMMENT		Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers 1. .2940 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:9330165116" /db_xref="MGI:2398424" /db_xref="taxon:10090" /clone="9330165116" /sex="male" /tissue_type="diencephalon" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 116. .2005 /notes="unnamed protein product; cholinergic receptor, nicotinic, alpha polypeptide 4 (MGD MGI:87888, GB NM_015730, evidence: BLASTN, 99%, match=1946) putative" /codon_start=1 /protein_id="BAC28638.1" /db_xref="GI:26329799" /translation="MSIGSGAPPPLLLPLLLLLGLTLLPASSHLETRAHERLLK RLPSGYNKSRPVANISDVVLRFGLSIAQLIDVDEKNQMTINWVKQEWHDYKLR DPGEYENISIRIPSELIRWPDVLYNNADGFVTHLTKAHLFYDGRVQWTPPIYK SSCSIDVTFPPDQKWTYKAFIRRLPFLFTINLIIPCLLSCTLVLFVFLVSECGEK TNYRKYECALPIPTYTAFIRRLPFLFTINLIIPCLLSCTLVLFVFLVSECGEK VTLCSVLSLTVFLLLITETIIPSTLVIPLIGELVLTMTVLTSLVITVFLVNHV RSPRTHMPAVRRVFLDIIVPLLFMKRPSVKDKRLLIESHKNANAPRFPFES EPGILGDI(CNGLSPATPCNRMDTAVETQPCRSFSHKVPDLTKSEVAKSPSPG SCFPNNSGAPLVKARSUSVQHPSSQEAEGSIRCSRISQYCVSODGAASLTESK PTGSPASLKRFLQPSQTSKCTCKECPSPVITVLKAGTKAPQHLPLSPAL TRAVEGVQVIADHLKAEDTDFSVKEDWKVYVMDIRFLWMEIIVCLLTGVLFLPFP LAGMI"
FEATURES		source
CDS		
ORIGIN		
Query Match	10.1%	Score 232.6; DB 11; Length 2940;
Best Local Similarity	56.6%	Pred. No. 2.4e-45;
Matches	452; Conservative	0; Mismatches 344; Indels 3; Gaps 1;
Qy	929	ATCATGAAAAGAGACTGTTACACGATCTTTTGGATCCTTATAATACACTAGAACGTCCCG 988
Db	225	ATGCGGAGGAGCGGCTCTCTGAAGAGACTCTTTCTCTGGCTACAAACAGTGGTCTCGGCCAG 284
Qy	989	TTCTCAATGAATCGGACCCGTTACAATTAAGCTTTGGTTAACTTTAATGCAAAATTATCG 1048
Db	285	TAGCCAATATCTCAGATGTGGTCTCTTGGCTTTGGCTTCTCGATTGCTCAGCTCATTTG 344
Qy	1049	ATGTGGACGAGAAAATCAATTGTGTAGTGCATTAATGTGTGTGTAACTGAGTGGGACG 1108
Db	345	ATGTGGATGAGAAAACCAAGATGATGACGACCAACAGTGTGGGTGAACAGAGGTGGCATG 404
Qy	1109	ACATGAATTCCTCGCTGGAACAACCTCGACTATGCGGAGTTAAGGATCTCGAAATACCGC 1168
Db	405	ACTACAAACTCGCTGGGACCCCTGGTGACTACGAGATGTCACTCCATCCGATCCCAT 464
Qy	1169	CGCATCGCATCTGGAAGCCGACGTGTGTATGATGATACACAGTGGCGGATGAGGATTTGACG 1228
Db	465	CTGAACCTCATCTGGAGGCTTGACATCGTCTCTACAACAAGCGGACGGGACTTTGCAG 524
Qy	1229	GCACCTACACAGAACGCTGTGTGGTGGGAAACAACGCTCGTGTCTATACGTTCCGCGG 1288
Db	525	TCACCCACCTAACCAAGAGCCCACTGTCTTATGATGGGCGTGTGCACTGGACACCCCGG 584
Qy	1289	GGATCTTCAAGTCGACGTGCAAGATCGACATCGATCGTGTTCCTCCCTTCGATGACACGCGT 1348
Db	585	CCATCTATAAGAGCTCTCGAGCATCGACGTCACTTCTTCCCTTCGACAGCAGAACT 644
Qy	1349	GCAGATGAAGTTGCGGAGTGGACCTACGACGATTCACGAGTTCACGCTGGATTTACAATTACAAG 1408
Db	645	GTACCATGAAGTTTGGGTCTTGACCTACCAAGCCCAAGATTGACTT---GGTGAGCA 701
Qy	1409	ATGAAACTGGCGGTGATATCAGCAGTTACGTGTCTCAACGGCGAGTGGGAACTACTCGGTG 1468
Db	702	TGCACAGCGGTGTGGACCAACTGCACTTCTGGGAAAGTGGGAGTGGGTCAATGTGGATG 761
Qy	1469	TGCGCGGCAACCTAAGAGATCTATTACAACCTGTGCGCCGGAACCCCTATATAGACATCA 1528
Db	762	CCGTGGGCACTACCAACACAGGAAGTATGATGCTGTGCGGAGATCTATCTCTGACATCA 821
Qy	1529	CTTCCGCACTATCATCCGCGGACGACACTGTACTATTCTTCAACCTGATCATACCTT 1588
Db	822	CTACGCGCTTATCATCCGCGGACTGCACTGTCTTCTACACCATCAACCTTATCATCCGT 881
Qy	1589	GTGTACTGATTGCTCCATGGCCTTGTCTCGGATTCACCTTCCGCGGAGATTCGGGTGAAA 1648
Db	882	GCCTGCTCATCTCTCGCTCACCGTGTGTCTTCTATCTGCGCTGGAGTGGCGGAGA 941
Qy	1649	AATTATCGGTGGTGTACCATCTTGTCTCGCTGACCGGTGTTCTTGAATATGTTGCGG 1708
Db	942	AGTCAAGCTGTGCATCTCGGTGCTGCTTCTCTCACCGGCTTCTCTCTCTCTCTCATCACCG 1001
Qy	1709	AGACAATCGCGGCTACTTC 1727
Db	1002	AGATCATCCGCTCCACTC 1020
RESULT 10		
LOCUS		AK083157 3230 bp mRNA linear HTC 20-SEP-2003
DEFINITION		Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched library, clone:C630019M18 product:cholinergic receptor, nicotinic, alpha polypeptide 4, full insert sequence.
ACCESSION		AK083157
VERSION		AK083157.1 GI:26350296
KEYWORDS		HTC; CAP trapper.
SOURCE		Mus musculus (house mouse)
ORGANISM		Mus musculus
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS		1 Carninci, P. and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
92279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3

REFERENCE
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Kigawa, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, F., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4

REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

REFERENCE
AUTHORS

The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3230)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://phantom.gsc.riken.go.jp/.

FEATURES
source

Location/Qualifiers
1. ..3230
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C630019M18"
/db_xref="MGI:2417536"
/db_xref="taxon:10090"

CDS

/clone="C630019M18"
/sex="male"
/tissue type="hippocampus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
36. .1925
/notes="unnamed protein product; cholinergic receptor,
nicotinic, alpha polypeptide 4 (MGD|MGI:87888,
GBINM.015730, evidence: BLASTN, 99%, match=1946)
putative"
/codon_start=1
/protein_id="BAC38788.1"
/db_xref="GI:26350297"

/translations="MEIGSGGAPPPPLLLPLLLLTGTLPPASSHIFRAHAERLLK
RLFGYNGKSRPVANISDVLVRFGLSIAQLIDVDEKNQMTTNNVWVQKQWHDKYLR
DPGDYENVTIRIPSELIWRPDIVLVYNADGFAVTHLTKAHLFYDGRVOWTVPALVYK
SSCSIDVTFFPQQCTMKFGSWYDKAKIDLVSMHSRVDQDQFWSGWEVIVDVG
TYNRKTECCAEIYFDITYAFIRRLPLFTINLIIFCLLSCLTIVLVEVLPSSCGEK
VTLCISVLLSLTFLLEITLIPSTLSLVIFLIGELTFTWIFVLSIVITVFLVNVHH
RSPRTHTMPAWRRVFLDIIVRLFMKRPVSKDNCRRLLIESMHKMANAPRPPEPES
EPGLGDCNQGLSPAPTFCNMDTAVETQTCRSPSHKVPDLKTSVEKASPCPSPG
SCHPPNSSGAPVLIKARLSVQHVPSQEAAGSIRCSRSIOYCVSQDGAASITSEK
PTGSPASLKRPSQLPVSQDOTSPCKCTCKBPSPVPTVLKAGTKAPPOHLPLSPAL
TRAVBGVQYIADHLKAEEDTFSVKEDKQYVMVIDRFLMWFIIVCLLGTGLFLPPW
LAGMI"

ORIGIN

Query Match	10.1%	Score 232.6;	DB 11;	Length 3230;
Best Local Similarity	56.6%	Pred. No. 2.5e-45;		
Matches 452;	Conservative 0;	Mismatches 344;	Indels 3;	Gaps 1;
Qy	929	ATCATGAAAAGAGACTGTTACAGATCTTTTGGATCCCTTATAATACACTAGAACGTCGCG	988	
Db	145	ATCGGAGGAGCGCTCTCTGAAGAGACTTCTCTGGCTACACAAGTGGTCTCGGCAG	204	
Qy	989	TTCTCATGAATCGGACCGGTTCAATTAAGCTTTGGTTTAACTTTAATGCAATTCG	1048	
Db	205	TAGCCAATATCTAGATGTGGTCTCTTGGCTTTGGCTTGTGCAATGCTCAGCTCAATG	264	
Qy	1049	ATGTGGACGAGAAAAAACAATTGCTAGTCACTAATGTGTGTTGTTAAAACTGGAGTGAACG	1108	
Db	265	ATGTGGATGAGAAAAACAGATGATGAGCACCACAGTGTGGTGAACGAGAGTGGCATG	324	
Qy	1109	ACATGAATCTCCGTGGAACACCTCCGACTATGGCGAGTTAAGATCTCGGAATACCGC	1168	
Db	325	ACTACAACTCGCTGGACCCCTGCTGACTACGAGATGTCCCTCCATCCGATCCCAT	384	
Qy	1169	CGCATCGCATCTGGAACCGGACGCTGCTGATGATACACAGTCCGATGAGGATTTGACG	1228	
Db	385	CTGAACCTCATCTGGAGGCTGACATCGTCTCTTACAAACGCGACGCGGACITTTGAG	444	
Qy	1229	GCACCTACACGAACTGGTGGTGGGAAACACGGCTCGTCTCTATACGTTCCGCGCG	1288	
Db	445	TCACCCACCTAACCAAGCCACCTGTTCTATGATGGCGGTGTCAGTGGACACCCCGG	504	
Qy	1289	GGATCTTCAAGTCGACGTCGAAGATCGACATCAGGTGTTCCCTTCGATGACGACGCGT	1348	
Db	505	CCATCTATAAGAGCTCTCTGCAGCTCGACGTCACGCTTCTCCCTTCGACGAGCAACT	564	
Qy	1349	CGGAGATGAAGTTCGGCAGTTGGACCTACGACGGATTCACGCTGGATTACAATTAACA	1408	
Db	565	GTACCATGAAGTTTGGTCTTGGACCTACGACAGGCCAAGATTGACTT---GGTGACA	621	
Qy	1409	ATGAACCTGGCGGTGATATCAGCAGTTACGTCCTCAACGGCGAGTGGGAATCTCTGGGTG	1468	
Db	622	TGCACAGCGGTGTGGACCACTGGACTTCTGGGAAAGTGGGAGTGGGTCTATTGTGGATG	681	
Qy	1469	TGCCCGGCAACGTTAAACAGATCTATTACAACCTGCTCCCGGACCCCTATATAGACATCA	1528	
Db	682	CCGTGGGCACCTACAAACACCGAGAGTATGAATGTGTGCCGAGATCTATCTCTGACATCA	741	
Qy	1529	CTTTCCGCTCATCATCCGCGGACGAACTGTACTATTTCTTCAACCTGTATCATACCTT	1588	

gene tries
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source Location/Qualifiers
1..1436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCM1369"
ORIGIN
Query Match 9.8%; Score 225.4; DB 29; Length 1436;
Best Local Similarity 56.1%; Pred. No. 1.1e-43;
Matches 446; Conservative 0; Mismatches 346; Indels 3; Gaps 1;
Qy 933 TGAAGAAGACTGTACAGATCTTTTGGATCCTTATAATACACTAGACGTCCCGTTCT 992
Db 20 TGAGCACCGCTCTATTGAGCGCGCTGTTTGAAGATTACAATGAGATCATCCGGCTGTGGC 79
Qy 993 CAATGAATCGGACCGCTTACAATTAAGCTTTGGTTTAACTTTAATGCAAAATATCGATGT 1052
Db 80 CAACGTGTCTGACCCAGTTCATCCATTTTCAGGTGTCCATGCTCAGCTGGTGAAGT 139
Qy 1053 GGACGAGAAAATCAATGTCTAGTCACTAATGTGTGGTTAAACCTGGAGTGGACGACAT 1112
Db 140 GGATGAAGTAAACAGATCATGTGAGACCAACCTGTGGCTCAAGCAAAATCTGGAATGACTA 199
Qy 1113 GAATCTCGCTGGACACCTCGACTATGCGGAGTAAAGATCTGGAATACCGCCGCA 1172
Db 200 CAAGCTGAAGTGAACCCCTCTGACTATGTTGGGGCAGAGTTCAFGCGTGTCCCTGACCA 259
Qy 1173 TCGCATCTGGAAGCGGACGTGCTGATACAAAGTGCAGTGGAGATTGACGGCAC 1232
Db 260 GAAGATCTGGAAGCAGACATTTGCTGTATACAACTGTTGGGATTTCCAGTGA 319
Qy 1233 CTACAGACGAACGTGGTGGTGGGAAACAACGGCTCGTGTCTATACGTTCCGCGGGGAT 1292
Db 320 CGACAAGACCAAGACCTTACTCAAGTACACTGGGAGGTGACTTGGATACCTCCGGCCAT 379
Qy 1293 CTTCAAGTCGAGTCGACAGATCGACATCACTGTTCCCTTCGATGACCGGTGCGA 1352
Db 380 CTTTAAAGAGTCCTGTAAATTCGAGTCGACCTACTTCCCGTTTGATTACCAAACTGAC 439
Qy 1353 GATGAAGTTCGGCAGTTCGACCTACGACGATTCAGCTGGATTTTACAATTAACAGATGA 1412
Db 440 CATGAAGTTCGGTTCCTGCTCTAGATAAGCGGAATCGAT---CTGGTCCCTGATCGG 496
Qy 1413 AACTGGCGGTGATATCAGAGTATAGTGTCTCAACGGCGAGTGGGAACCTACTGGGTGTGCC 1472
Db 497 CTCCTCCATGAACCTCAAGGACTATTGGGAGAGCGCGAGTGGGCCATCATCAAGCCCC 556
Qy 1473 CGGCAACGTAACGAGATCTATTACAACCTGCTGCCGGAACCTATATAGACATCACCTT 1532
Db 557 AGGCTACAAACAGACATCAGTACATCTGTCGAGGAGATCTACCCGACATCACATA 616
Qy 1533 CGCCATCATCATCCGCGGACGAACCTGTACTATTCTTCAACCTGATCATACCTTGTGT 1592
Db 617 CTCGCTGTATACCCGGCGCTCCCTTGTCTTACACCATCAACCTCATCATCCCTGCTGT 676
Qy 1593 ACTGATTCCTCATGGCCCTTCTCGGATTACCTGCGGCGGACAGATTCGGGTGAATAAT 1652

Db 677 GCTCATCTCTCTCTCTCACTGTGCTGCTTCTTACCTCCCTCCGACTCCGCTGGAAGGT 736
Qy 1653 ATCGCTGGGTGTTACCATCTTGTCTCGCTGACCGTGTCTTCTGAATATGTTGCCGAGAC 1712
Db 737 GACCTGTGATTTCTGCTCTCTCTCCCTGACGGTGTCTTCTCTCGTGAATCACTGAGAC 796
Qy 1713 AATGCCCGGCTACTTC 1727
Db 797 CATCCCTTCCACCTC 811
RESULT 13
AY402874 1436 bp DNA linear GSS 15-DEC-2003
LOCUS Pan troglodytes CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402874
VERSION AY402874.1 GI:39758857
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source Location/Qualifiers
1..1436
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCM1369"
ORIGIN
Query Match 9.6%; Score 222.2; DB 29; Length 1436;
Best Local Similarity 54.7%; Pred. No. 6.7e-43;
Matches 435; Conservative 0; Mismatches 357; Indels 3; Gaps 1;
Qy 933 TCAAAAGAGACTGTTTACAGATCTTTTGGATCCTTTATATACTAGTACGTCCCGTTCT 992
Db 20 TGAGCACCGCTCTATTGAGCGCGCTGTTTGAAGATTACAATGAGATCATCCGGCTGTGGC 79
Qy 993 CAATGAATCGGACCGCTTACAATTAAGCTTTGGTTTAACTTTAATGCAAAATATCGATGT 1052
Db 80 CAACGTGTCTGACCCAGTTCATCCATTTNNNGTGTCCATGCTNNNNTTGGTGAAGGT 139
Qy 1053 GGACGAGAAAATCAATGTCTAGTCACTAATGTGTGGTTAAACTGGAGTGGACGACAT 1112
Db 140 GGATGAAGTAAACAGATCATGGAGACCAACNTGNGCTCAAGCAAAATCTGGAATGACTA 199
Qy 1113 GAATCTCGCTGGAAACACTCCGACTATGGCGAGTATGAGATCTGCGAATACCGCCGCA 1172
Db 200 CAAGCTGAAGTGAACCCCTCTGACTATGTTGGGGCAGAGTTTCATCGGTGTCCCTGCNN 259
Qy 1173 TCGCATCTGGAAGCGGACGTGCTGATGTACAAAGTGGGATGAGGATTGACGGCAC 1232


```
Db      695 GGCAGAGAGTCACGCTGTGTCATCTCGGTGCTGCTTTCTCTCACCGTCTTCCTG 748

RESULT 15
LOCUS   AY406232
DEFINITION Mus musculus CHRNA1 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY406232
VERSION   AY406232.1 GI:39762206
KEYWORDS GSS.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1374)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
        Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
        Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
        Adams,M.D. and Cargill,M.
        Inferring nonneutral evolution from human-chimp-mouse orthologous
        gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1374)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
        Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
        Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
        Adams,M.D. and Cargill,M.
        Direct Submission
        Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
        Rockville, MD 20850, USA
TITLE   This sequence was made by sequencing genomic exons and ordering
        them based on alignment.
FEATURES             location/Qualifiers
source               1..1374
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10090"
                    <1..>1374
                    /gene="CHRNA1"
                    /locus_tag="HCM2488"
ORIGIN
Query Match          9.5%; Score 219.2; DB 29; Length 1374;
Best Local Similarity 53.8%; Pred. No. 3.5e-42;
Matches 452; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

Qy      904 TCTGCTAAAGTTTCCTAGCAGGATATCATGAAAGAGACTGTGTACAGATCTTTTGGAT 963
Db      40 TCCGCTGGCCTTGTCTCTGGGCTCCGAAACATGAGACGGCTCTGGTGGCAAGCTCTTTGAA 99

Qy      964 CTTTATATACACTAGACGTCCTGCTTCTCAATGAATCGGACCGCTTACATTAAGCTTT 1023
Db      100 GACTACAGCAGTGTAGTCGGCCAGTGGAGACCAACCGTGAGATTGTACAAGTCAACCGTG 159

Qy      1024 GGTTTAACTTTAAATGCAAAATTTATCGATGTGCAGCAGAAAAATCAATTGCTAGTCACTAAT 1083
Db      160 GGTCTACAGCTGATCCAGCTTATCAATGTGGATGAAGTAATCAGATTGTGACACCAAT 219

Qy      1084 GTGTGGTTAAACTGGAGTGAACGACATGAATCTCCGCTGGAAACACCTCCGACTATGGC 1143
Db      220 GTACGCTCTGAACAGCAATGGTTCGATTACAACCTTGAATGGAATCCAGATGACTATGGA 279

Qy      1144 GGAGTTAAGGATCTCGGAATACCGCGCATCGCATCTGGAAGCCGGAGCTGCTGATGTAC 1203
Db      280 GGAGTGAATAAAATTCACATCCCTCGGAAAAGATCTCGCGCCGGAGCTGCTCTCTAT 339

Qy      1204 AACAGTCCGGATGAGGGATTTTACCGGCACCTACCAGACGAACGAGTGGTGGTGGGAACAAC 1263
Db      340 AACAAACGACAGCGGCACTTTGCCATTGTCAAAATTCACCAAGGTGCTCTCTGGACTACAC 399

Qy      1264 GGCTGCTGTCTATACGTTCCGCGCGGGGATCTTTCAAGTCGACGTCGAAGATCGACATCAG 1323
```

```
Db      400 GGCACATCATCTGGACACCGCCAGCCATCTTTAAAAAGCTACTGTGAGATCATTTGTCACT 459
Qy      1324 TGGTTCCCTTTTCGATGACACCGGTGCGAGATGAAGTTTCGGCAGTTTCGACCTACGACGA 1383
Db      460 CACTTTCCCTTCGATGAGCAGAACTGAGCATGAAGCTGGCACCCTGGACCTATGACGGC 519
Qy      1384 TTCCAGCTGGAATTTACAAATTTACAAGATGAAAACTGGCGGTGATATCAGCAGTTACGTGCTC 1443
Db      520 TCTGTGTGGCCATTAAACCCGGAAGTGACACGCCGACCTGAGTAACCTTCATGGAGAGC 579
Qy      1444 AACGGCAGTGGGAACTACTGCGGTGTGCCCGGCAACGTAACGAGATCTATTACAATGC 1503
Db      580 GGGGAGTGGGTGATCAAGGAAGCTCGGGGCTGGAAGCACTGGGTGTTTCTACTCTCTGCTGC 639
Qy      1504 TGCCCGGAACCTTATATAGACATCACCTTCGCCCATCATCATCGCCGACGAACTGTAC 1563
Db      640 CCACCACTCCCTACCTGGACATCACCTACCTTGTGATGAGCGCCCTGCCCCCTCTAC 699
Qy      1564 TATTTCCTCAACCTGATCATACCTTGTGTACTGATTCCTCCATGCGCCTTGTCTCGGATTC 1623
Db      700 TTCATTGTCAACGTCATCATTCCTGCTGCTCTCTCTCTCTTCTTAACACAGCCTGGTGTTC 759
Qy      1624 ACCCTGCCGCGAGATTCGGGTGAAATTTATCGCTGGGTGTACCATCTTGTCTCTCGCTG 1683
Db      760 TACCTGCCACAGACTCAGGGGAGAAGATGACGCTGAGCATCTCTGTCTTACTGTCCCTG 819
Qy      1684 ACCGTGTTTCTGAATATGTTGCCGAGACAATGCGCGCTACTTCCGATGGGTGCCCATG 1743
Db      820 ACCGTGTTCTTCTGTCATTTGTGAGACTAATCCCTTCCACCTCCAGCGCTGTGCCCTG 879
```

Search completed: May 8, 2004, 06:38:48

Job time : 6225.93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:09:28 ; Search time 66.2366 Seconds
(without alignments)
3284.615 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKTEVDDELWLAV.....MFAIIATIVLLSAPHIIVS 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4043	100.0	770	3	AAY50814 D. melano
2	1850	45.8	496	3	AAY50815 H. viresc
3	1609	39.8	501	3	AAY50816 H. viresc
4	1498.5	37.1	327	4	ABB70382 Drosophil
5	1407.5	34.8	311	4	ABB63683 Drosophil
6	1179.5	29.2	502	2	AAW12368 Neuronal
7	1173.5	29.0	502	7	ADE57308 Rat Prote
8	1173.5	29.0	502	7	ADD47049 Rat Prote
9	1168	28.9	511	2	AAW12369 Neuronal
10	1159.5	28.7	502	2	AAW44153 Human neu
11	1159.5	28.7	502	2	AAW09025 Neuronal
12	1159.5	28.7	502	3	AAW24088 Human PRO
13	1159.5	28.7	502	3	AAW24088 Human PRO
14	1159.5	28.7	502	4	ABB2690 Nicotinic
15	1159.5	28.7	502	4	ABB50012 Wild-type
16	1159.5	28.7	502	5	ABB2435 Human neu
17	1159.5	28.7	502	7	ADA10874 Human neu
18	1159.5	28.7	502	7	ADA10874 Human neu
19	1159.5	28.7	502	7	ADA47051 Human PRO
20	1159.5	28.6	502	7	ADE57310 Human PRO
21	1153.5	28.5	502	4	AAW69216 V2747 var
22	1149.5	28.4	502	4	AAW50015 Mutant hu
23	1143.5	28.3	502	4	AAW50016 Mutant hu
24	1125	27.8	498	4	ABB50017 Mutant hu
25	989	24.5	193	4	ABB60432 Drosophil
					ABB60716 Drosophil

ALIGNMENTS

RESULT 1

AAAY50814

ID AAY50814 standard; protein; 770 AA.

XX AC AAY50814;

XX AC

DT 17-FEB-2000 (first entry)

XX DE

XX D. melanogaster acetyl-choline receptor protein from clone Da7.

KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; AChR.

XX OS

XX Drosophila melanogaster.

XX PN

XX DE19819829-A1.

XX PD

XX 11-NOV-1999.

XX PF

XX 04-MAY-1998; 98DE-01019829.

XX PR

XX 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.

XX PA

XX Aay50814 D. melano

XX Aay50815 H. viresc

XX Aay50816 H. viresc

XX ABB70382 Drosophil

XX ABB63683 Drosophil

XX AAW12368 Neuronal

XX ADE57308 Rat Prote

XX ADD47049 Rat Prote

XX AAW12369 Neuronal

XX AAW44153 Human neu

XX AAW09025 Neuronal

XX AAB24088 Human PRO

XX AAB2690 Nicotinic

XX AAB50012 Wild-type

XX ABB2435 Human neu

XX ABB70492 Human neu

XX ADA10874 Human neu

XX ADA10874 Human neu

XX ADA47051 Human PRO

XX ADE57310 Human PRO

XX AAW69216 V2747 var

XX AAW50015 Mutant hu

XX AAB50016 Mutant hu

XX ABB50017 Mutant hu

XX ABB60432 Drosophil

XX ABB60716 Drosophil

SQ Sequence 770 AA;

Aae12824 Caenorhab
Abp96318 Caenorhab
Aab50014 Chimeric
Aab50018 Mature ce
Aaw4155 Human neu
Aar73966 Alpha 2 s
Aaw09021 Neuronal
Abg61850 Prostate
Abg31800 Human neu
Abb82430 Human neu
Ada10855 Human neu
Adc71171 Human 205
Adc71169 Human 205
Add45584 Human Pro
Ade59169 Human Pro
Ade59167 Rat Prote
Add45583 Rat Prote
Abb59012 Drosophil
Ade57314 Human Pro
Ade57318 Human Pro

26 972 24.0 554 4 AAE12824
27 972 24.0 554 6 ABP96318
28 922.5 22.8 470 4 AAB50014
29 912 22.6 448 4 AAB50018
30 848.5 21.0 529 2 AAW4155
31 848.5 21.0 529 2 AAR73966
32 848.5 21.0 529 2 AAW09021
33 848.5 21.0 529 5 ABG61850
34 848.5 21.0 529 5 ABG31800
35 848.5 21.0 529 5 ABB82430
36 848.5 21.0 529 7 ADA10855
37 848.5 21.0 529 7 ADC71171
38 848.5 21.0 529 7 ADC71169
39 848.5 21.0 529 7 ADD45584
40 848.5 21.0 529 7 ADE59169
41 847.5 21.0 512 7 ADE59167
42 847.5 21.0 512 7 ADD45583
43 847.5 21.0 772 4 ABB59012
44 845 20.9 503 7 ADE57314
45 845 20.9 503 7 ADE57318

Query Match 100.0%; Score 4043; DB 3; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKNAQLKLTVEVDDDELAVLRLAHCSNFSSSSSTRTSSNQHNQQLTLQPRSLSTKH 60
 Db 1 MKNAQLKLTVEVDDDELAVLRLAHCSNFSSSSSTRTSSNQHNQQLTLQPRSLSTKH 60

Qy 61 HSNITASECHNSQQQEPASKEDVANHGSRNDQOHLQOLDSSNMLSPKTAATAAAGDEA 120
 Db 61 HSNITASECHNSQQQEPASKEDVANHGSRNDQOHLQOLDSSNMLSPKTAATAAAGDEA 120

Qy 121 TQOQTNIRLCARQRLRRRRKPKATENETDIKKQQQLSMPPEKTKRSTDTTSTAAT 180
 Db 121 TQOQTNIRLCARQRLRRRRKPKATENETDIKKQQQLSMPPEKTKRSTDTTSTAAT 180

Qy 181 TSCPTATMOCRASDNFSPISRHRDVRSTATFAWLHVQLVLSLQQWQLHVQORSVL 240
 Db 181 TSCPTATMOCRASDNFSPISRHRDVRSTATFAWLHVQLVLSLQQWQLHVQORSVL 240

Qy 241 LPRRIASTAFISYLGFAQLKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300
 Db 241 LPRRIASTAFISYLGFAQLKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300

Qy 301 LSAKVCLAGYHEKRLHLLDLPYNTLBRPVLNEDPLQSLFGLTLMQIIDVDEKNQLLVT 360
 Db 301 LSAKVCLAGYHEKRLHLLDLPYNTLBRPVLNEDPLQSLFGLTLMQIIDVDEKNQLLVT 360

Qy 361 NVWLKLEWMDNMLRWNTSDYGGVKDLRIPPHRIWKPVDVLYNSADEGFGTYQTNVVVRN 420
 Db 361 NVWLKLEWMDNMLRWNTSDYGGVKDLRIPPHRIWKPVDVLYNSADEGFGTYQTNVVVRN 420

Qy 421 NGSCLYVPPGIFKSTCKIDITWFFPDQORCEMKFGSWTYDGFQDLQLODETGDISYV 480
 Db 421 NGSCLYVPPGIFKSTCKIDITWFFPDQORCEMKFGSWTYDGFQDLQLODETGDISYV 480

Qy 481 LNWGEWELLGVPGKNEIYVCCPEYIDITFAIIIRRTLYFFNLIIPCVLIASMLLG 540
 Db 481 LNWGEWELLGVPGKNEIYVCCPEYIDITFAIIIRRTLYFFNLIIPCVLIASMLLG 540

Qy 541 FTLPDPSGEKLSLGVTLISLTVFLNVAETMPATSDAVPLWIRVFLCWLPIWLRMSRP 600
 Db 541 FTLPDPSGEKLSLGVTLISLTVFLNVAETMPATSDAVPLWIRVFLCWLPIWLRMSRP 600

Qy 601 GRPLILEFTPTPCSDTSSERKHQILSDVELKERSKSLANVLIDDDFRHNCRPMTPGG 660
 Db 601 GRPLILEFTPTPCSDTSSERKHQILSDVELKERSKSLANVLIDDDFRHNCRPMTPGG 660

Qy 661 TLPNPAFYRTVYGGDDSGIPGSTRMPDAVTHHTCIKSTVEYELGLILKEIRFTDQ 720
 Db 661 TLPNPAFYRTVYGGDDSGIPGSTRMPDAVTHHTCIKSTVEYELGLILKEIRFTDQ 720

Qy 721 LRKDECDNDIANDKFAAMVVDRLCLIFTMFAIATIAVLLSAPHIIVS 770
 Db 721 LRKDECDNDIANDKFAAMVVDRLCLIFTMFAIATIAVLLSAPHIIVS 770

RESULT 2
 ID AAY50815
 AC AAY50815; standard; protein; 496 AA.
 AC AAY50815;
 XX 17-FEB-2000 (first entry)
 DE H. virescens acetyl-choline receptor protein from clone Hva7-1.
 KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 OS neurotransmission; plant protection agent; conductance; AChR.
 XX Heliothis virescens.
 XX

DE19819829-A1.
 11-NOV-1999.
 04-MAY-1998; 98DE-01019829.
 04-MAY-1998; 98DE-01019829.
 (FARB) BAYER AG.
 Adamczewski M, Oellers N, Schulte T;
 WPI; 2000-014207/02.
 N-PSDB; AAZ24476.
 New nucleic acid encoding a nicotinic acetylcholine receptor from insects, used to identify potential insecticides.
 Example 1a; Page 17-19; 26pp; German.
 This invention describes a novel nucleic acid (NA) encoding a nicotinic acetyl-choline receptor (I) from insects which can be used as an insecticide. Inhibitors of (I) interfere with neurotransmission. (I) (also vectors containing it, its regulatory regions, and antibodies directed against (I)-encoded proteins) are used to screen for: (a) plant protection agents that alter conductance of AChR, potentially useful as insecticides, or (b) genes which encode polypeptides that are involved in formation of functionally related AChR in insects. (I) are also used to isolate and characterize the specified regulatory regions and for recombinant production of (II). This sequence represents an acetyl-choline receptor isolated from *Heliothis virescens*

Query Match 45.8%; Score 1850; DB 3; Length 496;
 Best Local Similarity 70.0%; Pred. No. 7.2e-161;
 Matches 361; Conservative 30; Mismatches 49; Indels 76; Gaps 7;

Qy 295 LLIYVNLISAKVCLAGYHEKRLHLLDLPYNTLBRPVLNEDPLQSLFGLTLMQIIDVDEK 354
 Db 17 LLLCCLLWPGARGCYHEKRLHLLDLPYNTLBRPVLNEDPLQSLFGLTLMQIIDVDEK 76

Qy 355 NQLLVTVNLKLEWMDNMLRWNTSDYGGVKDLRIPPHRIWKPVDVLYNSADEGFGTYQ 414
 Db 77 NQLLVTVNLKLEWMDNMLRWNTSDYGGVKDLRIPPHRIWKPVDVLYNSADEGFGTYQ 136

Qy 415 NVVVRNNGSCLYVPPGIFKSTCKIDITWFFPDQORCEMKFGSWTYDGFQDLQLODETG 474
 Db 137 NVVVRNNGSCLYVPPGIFKSTCKIDITWFFPDQORCEMKFGSWTYDGFQDLQLODETG 196

Qy 475 DISSVVLNGEWELLGVPGKNEIYVCCPEYIDITFAIIIRRTLYFFNLIIPCVLIA 534
 Db 197 DISSVVLNGEWELLGVPGKNEIYVCCPEYIDITFAVIRKRLTYFFNLIIPCVLIA 256

Qy 535 SMALLGFTLPDPSGEKLSLGVTLISLTVFLNVAETMPATSDAVPL----- 581
 Db 257 SMALLGFTLPDPSGEKLSLGVTLISLTVFLNVAETMPATSDAVPLGTYFNCIMFVA 316

Qy 582 -----WIRVFLCWLPIWLRMSRP---RPLIERPTPCS 614
 Db 317 SSVVSTILNLNHYHRHADTHMSDWIRCVFLWFLWLRMSRPGSATPPPARVPPDP-- 374

Qy 615 DTSSERKHQILSDVELKERSKSLANVLIDDDFRHNCRPMTPGGTLPNPAFYRTVY 674
 Db 375 -----DLERERSKSLANVLIDDDFRH-----PQAQPPQCCRYR---- 412

Qy 675 QGDGSGIPGSTRMPDAVTHHTCIKSTVEYELGLILKEIRFTDQLRKDECDNDIANDW 734
 Db 413 GGEENGAG-----LAAHSCF--GVDEVLSLILKEIRVITDQMKRDEDEDADISRDW 460

Qy 735 KFAAMVVDRLCLIFTMFAIATIAVLLSAPHIIVS 770
 Db 461 KFAAMVVDRLCLIFTMFAIATIAVLLSAPHIIVS 496

Qy	582		-----WIRIVFLCWLPMILMSRPGRLILEFTTFCSDTS	61
Dd	306	SSVVLTVLVNHHRTADIHMPWKISFLOLWPLMLSPGKIKTK---	TTMMNR	362
Qy	618	SERKHQLSDVELKERSKSLANVLIDDDFRHNCRPMTG----	GTLPHNPAYRTV	672
Dd	363	-----MRELELKERSKSLANVLIDDDFRHG--PPPNSTASTGNLGPGCSIFRTD	413	
Qy	673	YGOG-----DDGSIGPIGSTRMEDAVTHETCIKSSTSEYELGLILKEIREITDQLRKD	724	
Dd	414	FRSFRPSTMEDVG--GGIGS-----HR-----EHLLRELQFITARMKKA	455	
Qy	725	DECNDIANDKFAAMVDRLCLIIITFMFAILATIAVLISAPHIIV	769	
Dd	456	DEEALISDWKFAAMVVDRFCLVFVLTFTTIATVAVLISAPHIIV	500	
RESULT 4				
ID	ABB70382			
XX	AC	ABB70382 standard; protein; 327 AA.		
XX	DT	26-MAR-2002 (first entry)		
XX	DE	Drosophila melanogaster polypeptide SEQ ID NO 37938.		
XX	KW	Drosophila; developmental biology; cell signalling; insecticide;		
XX	OS	Drosophila melanogaster.		
XX	PN	WO200171042-A2.		
XX	PD	27-SEP-2001.		
XX	PF	23-MAR-2001; 2001WO-US009231.		
XX	PR	23-MAR-2000; 2000US-0191637P.		
XX	PR	11-JUL-2000; 2000US-00614150.		
XX	PA	(PEKE) PE CORP NY.		
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;		
XX	XX	WPI; 2001-656860/75.		
XX	DR	N-PSDB; ABL14485.		
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.		
XX	PS	Disclosure; SEQ ID NO 37938; 2lpp + Sequence Listing; English.		
XX	CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU16176-ABU16051), expressed DNA sequences (ABU16176-ABU16175) and the encoded proteins (ABBS7737-ABBY2072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences		
XX	SQ	Sequence 327 AA;		
Query Match				
Best Local Similarity 37.1%; Score 1498.5; DB 4; Length 327;				
Matches 297; Conservative 4; Mismatches 1; Gaps 1;				

QY 1 MKNAQLKLTVDDELWLAHLACSNFSSSSSTRTSSNQHRNQQLTTLQPSLSSTKH 60
Db 1 MKNAQLKLTVDDELWLAHLACSNFSSSSSTRTSSNQHRNQQLTTLQPSLSSTKH 60
QY 61 HSNIASQHNSSQOEPAKDEVDVANHGSRNDQOHLQQLDSSNMLSPKTAATAAGDEA 120
Db 61 HSNIASQHNSSQOEPAKDEVDVANHGSRNDQOHLQQLDSSNMLSPKTAATAAGDEA 120
QY 121 TQOPTNIRLCARQKRLRRRKPKATPNETDIKKOQQLSMPFPKTRKSTDTYSTPAAT 180
Db 121 TQOPTNIRLCARQKRLRRRKPKATPNETDIKKOQQLSMPFPKTRKSTDTYSTPAAT 180
QY 181 TSCTATYMCRASDNFSPISRHDRVSTATPAWLHVQLVLSLQOQLHVQOQSVL 240
Db 181 TSCTATYMCRASDNFSPISRHDRVSTATPAWLHVQLVLSLQOQLHVQOQSVL 240
QY 241 LFRRIAASTIAFISYLGFAAQLKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 299
Db 241 LFRRIAASTIAFISYLGFAAQLKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300
QY 300 NLSAK 304
Db 301 NLSAK 305

RESULT 5
ABB63683
ID ABB63683 standard; protein; 311 AA.
XX AC ABB63683;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 17841.
XX DE Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX KW Drosophila melanogaster.
XX OS
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABL07786.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Disclosure; SEQ ID NO 17841; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 311 AA;
Query Match 34.8%; Score 1407.5; DB 4; Length 311;
Best Local Similarity 88.6%; Pred. No. 1.8e-120;
Matches 257; Conservative 18; Mismatches 12; Indels 3; Gaps 1;
QY 295 LLIV--LNLAKVCLAGYHEKRLHLLDLPYNTLFRPVLNESPQLQSLFGLTLMQIIDV 351
Db 19 MLVYGLLLTMIACAGPHEKRLHLLDLPYNTLFRPVLNESPQLQSLFGLTLMQIIDV 78
QY 352 DEKNQLLVTVWMLKLEWMDNMLRWNTSDYGVKDLRIPPHRIWKPVDVLMNSADEGFDGT 411
Db 79 DEKNQLLVTVWMLKLEWMDNMLRWNTSDYGVKDLRIPPHRIWKPVDVLMNSADEGFDGT 138
QY 412 YQTNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDORCEMKFGSWTYDGFQDLQLODE 471
Db 139 YATNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDORCEMKFGSWTYDGFQDLQLODE 198
QY 472 TGGDISSVVLNGEWELLGVFGKNEIYVNCPEPYIDITFAIIRRTLYFFFNLIIPCV 531
Db 199 AGGDISSFITNGEWDLLGVFGKNEIYVNCPEPYIDITFAIIRRTLYFFFNLIIPCV 258
QY 532 LIASMAILGFTLPDPSGEKLSLGVTTILLSTVFLNMVAETMPATSDAVPL 581
Db 259 LIASMAILGFTLPDPSGEKLSLGVTTILLSTVFLNMVAETMPATSDAVPL 308
RESULT 6
AAW12368
ID AAW12368 standard; protein; 502 AA.
XX AC AAW12368;
XX DT 17-JUN-1997 (first entry)
XX DE Neuronal alpha-bungarotoxin binding protein alpha subunit.
XX KW Neuronal alpha-bungarotoxin binding protein alpha 1; cholinergic;
XX KW ligand binding; ion channel.
XX OS Gallus sp.
XX FH Key Location/Qualifiers
XX FT Peptide 1..22 /label= Sig_peptide
XX FT Protein 23..502 /label= Mat_protein
XX FT US5599709-A.
XX PN 04-FEB-1997.
XX PD 28-SEP-1989; 89US-00413947.
XX PR 28-SEP-1989; 89US-00413947.
XX PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX PI Lindstrom JM, Schoepfer RD;
XX DR WPI; 1997-118297/11.
XX DR N-PSDB; AAT59196.
XX PT New isolated neuronal alpha-bungarotoxin-binding protein DNA - used to
XX PT screen cholinergic agents and other drugs which may affect ligand
XX PT binding, ion channel or other activities of the protein.
XX PS Example; Fig 2A-B; 18pp; English.
XX CC The alpha subunit (AAW12368) and alpha2 subunit (AAW12369) amino acid
XX CC sequences of chick neuronal alpha-bungarotoxin binding protein (ABBP)
XX CC were deduced from newly isolated DNA molecules (AAT59196-97) obcd. from
XX CC an 18-day chick embryo cdna library. ABBP subunits can be produced in

CC The present sequence is that of the alpha subunit of human nicotinic
 CC acetylcholine receptor (nAChR). The sequence includes regions that are
 CC conserved throughout the various nAChR alpha subunits and which are
 CC essential for ligand binding. The invention relates to water-soluble
 CC ligand-binding proteins derived from molluscs, especially acetylcholine-
 CC binding proteins (AChBPs) and analogues of ligand-gated ion channels,
 CC their crystals, and their use for screening ligands of ligand-gated ion
 CC channels. The water-soluble ligand-binding proteins are capable of
 CC forming multimers and are amenable to crystallization. The crystal
 CC structure of AChBP is provided, and can be used to generate 3D models of
 CC the extracellular ligand-binding domain of ligand-gated ion channels and
 CC thus for screening of drugs that act on these ion channels. Chimeric
 CC proteins are provided that are capable of binding a ligand of a ligand-
 CC gated receptor, and comprise at least the amino acids of the AChBP
 CC determining solubility of the AChBP, in the same positions as in the
 CC AChBP, and also comprising amino acids determining binding to the ligand.
 CC In the chimeric proteins, at least the essential amino acids of at least
 CC 1 of the conserved regions of an nAChR have been substituted for the
 CC corresponding amino acids, and preferably entire stretches have been
 CC substituted. New drugs can be developed that selectively intervene in
 CC neuronal signalling pathways, especially where the ligand-gated ion
 CC channel is the nAChR, and the related disorder is Tourette's syndrome,
 CC Alzheimer's disease, addiction to nicotine or schizophrenia
 XX

SQ Sequence 502 AA;

Query Match 28.7%; Score 1159.5; DB 4; Length 502;
 Best Local Similarity 45.1%; Pred. No. 2.6e-97;
 Matches 237; Conservative 82; Mismatches 115; Indels 91; Gaps 12;

QY 297 IYLNLSAK---VCLAGYHEKRLHLLDPYNTLPRVNLSPDLQSLFGLTMOIIVDE 353
 DB 8 VWLALASLLHVSLOGEFQKLYKELVKNYPLERPVANDSQPLTVVPSLSILQIMDVDE 67
 QY 354 KNQLLVTVNWLKLEWDMNLWNTSDYGVKDLRIPPHRIWKPDPVLMVNSADEGFGDTYQ 413
 DB 68 KNQVLTVNWLKLEWDMNLWNTSDYGVKDLRIPPHRIWKPDPVLMVNSADEGFGDTYQ 127
 QY 414 TNVVRNNGSLVYPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGFQDLQLOQDET 473
 DB 128 TNVLNNGSLVYPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGFQDLQLOQDET 184
 QY 474 GDISSYVLNGEWELLGVPGKNEIYVNCPCPEYIDITFAIIRRTLYFFNLIIPCVLI 533
 DB 185 ADISGYIPNGEWDLVGIPGKSERFECCKEYPDPVTFVTMRRTLYYGLNLIIPCVLI 244
 QY 534 ASMLLGLTLPDSDGKLSGLVTLLSLTVFLNVAETMPATSDAVL----- 581
 DB 245 SALALLVFLPADSGEKISLIGITVLLSLTVFLNVAETMPATSDAVL----- 612
 QY 582 -----WIRIVFLCWLPIRLMSRPG-----RPLILEFFPTT 612
 DB 305 GLSVVTVTVIVLYHHDDPGCKMPKWTIVILLNWCWFLMKRPGEDKVRP-AQCHKQR 363
 QY 613 CSDTSSERKHQILSDVELKERSKSLANVLIDDDFRH-----NCRPMTPGTTLPHNPAF 668
 DB 364 CSLASVE-----MSAVAPPPASNGNLLY-----IGFGLDGVHCV-TP----- 401
 QY 669 YRTVYGGDDGSGIGPIGSTRMPDAVTH-----HTCIKSTVEYELGLILKEIRFTDQLRK 723
 DB 402 -----DSGV-----CGRWACGPTHDEHLHGGQPPGDPDLAKILEEVRYIANRRC 449
 QY 724 DDECDIANDWKFAAMVDRICLIIFTWTFAILATIIVLLSAPHII 768
 DB 450 QDESEAVCSKFAACVVDRLCLMAFSVFTIICITIGILMSAPNFV 494

RESULT 14

AAB50012

ID AAB50012 standard; protein; 502 AA.

XX

AC AAB50012;

XX

DT 14-MAR-2001 (first entry)
 XX
 DE Wild-type human alpha7 ligand gated ion channel.
 XX
 KW Human; alpha7 nicotinic acetylcholine gated ion channel;
 KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance.
 XX
 OS Homo sapiens.
 XX
 PN WO2000073431-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 25-MAY-2000; 2000WO-US011862.
 XX
 PR 27-MAY-1999; 99US-0136174P.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Groppi VE, Wolfe ML, Berkenpas MB;
 XX
 DR WPI; 2001-061524/07.
 DR N-PSDB; AAC90380.
 XX
 PT Special cell culture medium for treating cells and for inducing mammalian
 PT cell lines to conduct calcium ions, comprising specified concentrations
 PT of ions of sodium, calcium and potassium at specified pH.
 XX
 PS Disclosure; Page 61-63; 77pp; English.
 XX

CC The present sequence is wild-type human alpha7 nicotinic acetylcholine
 CC gated ion channel. The human alpha7 ion channel was used in the
 CC construction of an alpha7/5-hydroxytryptamine (5-HT3) chimeric ligand
 CC gated ion channel (see AAC90382 and AAB50014). The alpha7/5-HT3 chimeric
 CC ion channel can be expressed by recombinant cells in the present
 CC invention, resulting in preferential calcium ion conductance by the cells
 XX

SQ Sequence 502 AA;

Query Match 28.7%; Score 1159.5; DB 4; Length 502;
 Best Local Similarity 45.1%; Pred. No. 2.6e-97;
 Matches 237; Conservative 82; Mismatches 115; Indels 91; Gaps 12;

QY 297 IYLNLSAK---VCLAGYHEKRLHLLDPYNTLPRVNLSPDLQSLFGLTMOIIVDE 353
 DB 8 VWLALASLLHVSLOGEFQKLYKELVKNYPLERPVANDSQPLTVVPSLSILQIMDVDE 67
 QY 354 KNQLLVTVNWLKLEWDMNLWNTSDYGVKDLRIPPHRIWKPDPVLMVNSADEGFGDTYQ 413
 DB 68 KNQVLTVNWLKLEWDMNLWNTSDYGVKDLRIPPHRIWKPDPVLMVNSADEGFGDTYQ 127
 QY 414 TNVVRNNGSLVYPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGFQDLQLOQDET 473
 DB 128 TNVLNNGSLVYPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGFQDLQLOQDET 184
 QY 474 GDISSYVLNGEWELLGVPGKNEIYVNCPCPEYIDITFAIIRRTLYFFNLIIPCVLI 533
 DB 185 ADISGYIPNGEWDLVGIPGKSERFECCKEYPDPVTFVTMRRTLYYGLNLIIPCVLI 244
 QY 534 ASMLLGLTLPDSDGKLSGLVTLLSLTVFLNVAETMPATSDAVL----- 581
 DB 245 SALALLVFLPADSGEKISLIGITVLLSLTVFLNVAETMPATSDAVL----- 612
 QY 582 -----WIRIVFLCWLPIRLMSRPG-----RPLILEFFPTT 612
 DB 305 GLSVVTVTVIVLYHHDDPGCKMPKWTIVILLNWCWFLMKRPGEDKVRP-AQCHKQR 363
 QY 613 CSDTSSERKHQILSDVELKERSKSLANVLIDDDFRH-----NCRPMTPGTTLPHNPAF 668
 DB 364 CSLASVE-----MSAVAPPPASNGNLLY-----IGFGLDGVHCV-TP----- 401
 QY 669 YRTVYGGDDGSGIGPIGSTRMPDAVTH-----HTCIKSTVEYELGLILKEIRFTDQLRK 723

Db 402 -----DSGV-----CGRMACSPTHDEHLHGGQPPEGDPDLAKILEEVRYIANRFR 449

Qy 724 DDECNDIANWKFAAMVVDRLCLIIFTMFAILATIAVLLSAPHII 768

Db 450 QDESEAVCEWKFACVVDRLCLMAFSVFTICTIGILMSAPNEV 494

RESULT 15

ABG70492

ID ABG70492 standard; protein; 502 AA.

XX AC ABG70492;

XX DT 06-DEC-2002 (first entry)

XX DE Human neuronal nicotinic acetylcholine receptor alpha 7 subunit.

XX KW Human; neuronal nicotinic acetylcholine receptor; nNACHr; receptor;

XX KW ion flux; alpha 7 subunit.

XX OS Homo sapiens.

XX PN US6440681-B1.

XX PD 27-AUG-2002.

XX PF 07-JUN-1995; 95US-00487596.

XX PR 03-APR-1990; 90US-00504455.

XX PR 30-NOV-1992; 92US-00938154.

XX PR 08-MAR-1993; 93US-00028031.

XX PR 08-NOV-1993; 93US-00149503.

XX PA (MERI) MERCK & CO INC.

XX PI Elliott KJ, Ellis SB, Harpold MM;

XX WPI: 2002-711528/77.

DR N-PSDB; ABS54875.

XX

PT Identifying antagonists or agonists of human neuronal nicotinic

PT acetylcholine receptors, by contacting recombinant cells with test

PT compound, and measuring ion flux of cells or binding of compound to

PT nNACHr.

XX

PS Claim 101; Col 59-84; 56pp; English.

CC The invention relates to a method for identifying compounds that are

CC antagonists or agonists of human neuronal nicotinic acetylcholine

CC receptors (nNACHrs), by contacting recombinant cells with a test compound

CC and measuring ion flux, the electrophysiological response of the cells or

CC binding of the test compound to the nNACHr. The recombinant cells are

CC produced by transfection with a nucleic acid encoding at least one human

CC nNACHr (alpha or beta) subunit, such that the cells express an nNACHr

CC comprising one human subunit encoded by the transfected nucleic acid.

CC This sequence represents the alpha 7 subunit of the human nNACHr

XX polypeptide

XX

SQ Sequence 502 AA;

Query Match 28.7%; Score 1159.5; DB 5; Length 502;

Best Local Similarity 45.1%; Pred. No. 2.6e-97;

Matches 237; Conservative 82; Mismatches 115; Indels 91; Gaps 12;

Qy 297 IYLNLSAK---VCLAGYHEKRLHLDLPYNTLPRPVLNESDPLQLSFGLTLMQIIDVDE 353

Db 8 VMLALAASLLHVSQGEFQKLYKELVKNYPLERPVANDSQPLTVVYFSLQLQIMDVDE 67

Qy 354 KNQLLVITNVMKLEWDMNLRWNTSDYGGVKDLRIPPHRIWKPDVLNYSADGFDCTYQ 413

Db 68 KNQVLATTNIWLQMSWTDHYLQWNVSEYPGVKTVAFPDGQIWKPDILLYNSADERFDATFH 127

Qy 414 TNNVVRNNGSLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFOLDLQLODETG 473

Db 128 TNVLVNSSGHCQYLPPGIFKSSCYIDVRWPFDFVQHCKLXPGSWSYGGWSLDLQMQE--- 184

Qy 474 GDISSYVLNGEWELLGVGPKRNEIYNYCCPEPYDITFAIIIRRTLYYFFENLIIPCVLI 533

Db 185 ADISGYIPNGEWDLVGIPGKRSEFYECCKEYPDVTFTVTRRTLYYGLNLLIPCVLI 244

Qy 534 ASMALGFTLPPDSGKLSLGVITLLSLTVFLNVAETMPATSDAVEL----- 581

Db 245 SALALLVFLPADSGEKISLGITVLLSLTVFMLIVAEIMPATSDSVELIAQYFASMTIIV 304

Qy 582 -----WIRIVFLQWLPLRMSREG---RPLILEFFPTP 612

Db 305 GLSVVVTVIVLQYHHHDPDGGKMPKTRVILLNWCWFLMKRPGEDKVRP-ACQHKQRR 363

Qy 613 CSDTSERKHQILSDVELKERSKSLANVLIDDDFRH---NCRPMTFGGTLPHNPAF 668

Db 364 CSLASVE-----MSAVAPPPASNGNLLY-----IGFGLDGVHCV-TP----- 401

Qy 669 YRTVYGGDDGSGTGPSTGTRMPDAVTH-----HTCIKSSTEYELGLLKEIRFTDOLRK 723

Db 402 -----DSGV-----CGRMACSPTHDEHLHGGQPPEGDPDLAKILEEVRYIANRFR 449

Qy 724 DDECNDIANWKFAAMVVDRLCLIIFTMFAILATIAVLLSAPHII 768

Db 450 QDESEAVCEWKFACVVDRLCLMAFSVFTICTIGILMSAPNEV 494

Search completed: May 7, 2004, 11:35:46

Job time : 69.2366 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:45:49 ; Search time 8474.79 Seconds
(without alignments)
3938.047 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKLTVDDELMLAV.....MEALLATIAVLISAPHIIVS 770

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xln

-Q=/cgn2_1/USPTO.spool/US09303232/runat_07052004_101110_23883/app_query.fasta_1.2261

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZB=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09303232@cgn_1_1_8225@runat_07052004_101110_23883 -NCPU=6 -ICPU=3

-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4043	100.0	2886	6	E58346 Nucleic aci
2	4043	100.0	2886	6	AX009610 Sequence
3	4014.5	99.3	2907	3	AF272778 Drosophil
4	3804.5	94.1	2834	3	AY036613 Drosophil
5	1991	49.2	1683	3	DME554210
6	1850	45.8	3629	3	AF143846 Heliothis
7	1850	45.8	3700	6	AX009612 Sequence
8	1850	45.8	3701	6	E58347 Nucleic aci
9	1625.5	40.2	2023	3	AF321446 Drosophil
10	1622.5	40.1	2023	3	AF321445 Drosophil
11	1621.5	40.1	1665	3	DME554209
12	1621.5	40.1	1699	3	BT011147 Drosophil
13	1620.5	40.1	2023	3	AF321447 Drosophil
14	1611	39.8	2110	3	AF321449 Drosophil
15	1609	39.8	3029	3	AF143847 Heliothis
16	1609	39.8	3109	6	E58348 Nucleic aci
17	1609	39.8	3109	6	AX009614 Sequence
18	1602	39.6	2068	3	AF321448 Drosophil
19	1567.5	38.8	2116	3	AY036614 Drosophil
20	1456.5	36.0	42079	2	AC017620 Drosophil
21	1456.5	36.0	158758	3	AC092243 Drosophil
22	1456.5	36.0	186803	3	AC092225 Drosophil
23	1456.5	36.0	272521	3	AE003642 Drosophil
24	1456.5	36.0	320754	3	DROSADH02
25	1199	29.7	1719	5	AY247962 Danio rer
26	1185.5	29.3	2088	10	AF225980 Mus muscu
27	1183.5	29.3	2106	10	RATNARAD L31619 Rattus ratt
28	1182.5	29.2	1848	10	MUSNARS L37663 Mus musculu
29	1179.5	29.2	2037	5	GGNNARA7A X68586 G.gallus mR
30	1179.5	29.2	2696	5	GGY7NAREC X52295 Chicken alp
31	1174	29.0	1542	5	AY299466 Takifugu
32	1173.5	29.0	3030	10	S53987 nicotinic r
33	1170.5	29.0	2090	5	GGANAREC X52296 Chicken alp
34	1163	28.8	1668	5	AY298752 Takifugu
35	1160.5	28.7	1551	4	BTAT X93604 B.taurus mR
36	1160.5	28.7	2087	9	HSARA7A X70297 H.sapiens m
37	1159.5	28.7	1509	9	AX054567 Sequence
38	1159.5	28.7	1509	9	HSNACHRA7 Y08420 H.sapiens m
39	1159.5	28.7	1876	6	AR055255 Sequence
40	1159.5	28.7	1876	6	AR071403 Sequence
41	1159.5	28.7	1876	6	AR173187 Sequence
42	1159.5	28.7	1876	6	AR224030 Sequence
43	1159.5	28.7	1876	6	AR282833 Sequence
44	1159.5	28.7	1876	6	AX19088 Sequence
45	1159.5	28.7	1876	9	HSU62436 Human nicot

ALIGNMENTS

RESULT 1

```

E58346
LOCUS       E58346             2886 bp    DNA        linear    PAT 18-JUN-2001
DEFINITION   Nucleic acid encoding insect acetyl choline receptor subunit.
ACCESSION   E58346
VERSION      E58346.1  GI:13019345
KEYWORDS     JP 2000023680-A/1.
SOURCE       Drosophila melanogaster (fruit fly)
ORGANISM     Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
              1 (bases 1 to 2886)
REFERENCE    Martin, A., Nadja, E. and Thomas, S.
AUTHORS      Nucleic acid encoding insect acetyl choline receptor subunit
TITLE        Patent: JP 2000023680-A 1 25-JAN-2000;
JOURNAL      BAYER AG
COMMENT      OS Drosophila melanogaster
              PN JP 2000023680-A/1
              PD 25-JAN-2000
              PF 26-APR-1999 JP 1999118159
              PR 04-MAY-1998 DE 19819829.9
              PI MARTIN ADAMUTSUBESUKI, NADJA ERASU, THOMAS SCHULTE PC
              C12N15/09, A01K67/033, C07K14/705, C07K16/28, C12N1/21, C12N5/10, PC
              C12Q1/68,
              PC G01N33/15, G01N33/50// (C12N1/21, C12R1:19), C12N15/00, C12N5/00 CC

FEATURES             Location/Qualifiers
     FT             CDS
     source          1..2886
                    /organism="Drosophila melanogaster"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:7227"

ORIGIN
Alignment Scores:
Pred. No.:          7,88e-286      Length:      2886
Score:              4043.00         Matches:    770
Percent Similarity: 100.00%         Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:        100.00%         Indels:    0
DB:                  6              Gaps:      0

US-09-303-232-2 (1-770) x E58346 (1-2886)

Qy      1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
Db      372 ATGAAAATGCACAACTGAACTGCACTGAACTTGACGATGATGAGCTGTGGCTGGCAGTA 431
Qy      21 ArgLeuAlaHisCysSerSerAsnPheSerSerSerSerSerSerSerSerSerSerSer 40
Db      432 AGATTAGGCGACTGTCAGCAGCACTTTAGCAGCAGTAGCAGCAGCAAGAACCAACGACGACG 491
Qy      41 AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
Db      492 AACGAGGCGCACACCACTGCAACACTGCAACCTGCAACCAAGGAGCTTAAGTACAAAACAC 551
Qy      61 HisSerAsnIleAlaSerGluGlnHisAsnSerGlnGlnGlnGluProAlaSerLysAsp 80
Db      552 CACAGCAACATTGCAAGCGAGCAGCACCAATAGCCAGCAACAGGAGCCAGCATCGAAGGAC 611
Qy      81 GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp 100
Db      612 GAGGATGTAGCAACACCGTAGAAGCAATGACCCAGCAGAGCGCATCTGCAACAGCTAGAC 671
Qy      101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla 120
Db      672 AGCAGCAACATGTTGTCGCAAAAGACAGCCGCGAGCAAACTGCTGCCGCGGCGATGAAGCA 731
Qy      121 ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
Db      732 ACAACCCCAACACCAACCAACATAAGACTGTGTGCACGACGACGACCAACGATTGCGTCGC 791

```

```

Qy      141 ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLysLysGlnGlnLeu 160
Db      792 CGACGAAAAAGAAAAACCGCAACCCCAACGAAACAGATATCAAGAAACACACGCAACTT 851
Qy      161 SerMetProProPheLysThrArgLysSerThrAspThrTyrSerThrProAlaAlaThr 180
Db      852 AGCATGCTCCCTTCAAAACCGGCAAAATCCACGGACACCTACAGCACACACGACGCAACA 911
Qy      181 ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGluPheSerIle 200
Db      912 ACCAGCTGTCGACAGCACCTCATCATGTCAGCCAGCGCAATAGTTCAGTATT 971
Qy      201 ProfileSerArgHisAspArgValSerThrAlaThrPheAlaTrpValLeuHisValLeu 220
Db      972 CCGATATCGAGACATGATAGATATCCAGGCCACATTCGCTGGGTGTTGCATGTGCTG 1031
Qy      221 GlnValLeuLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu 240
Db      1032 CAGGTGCTGCTGCTGCTGCTGCAACAGTGGCAACTTCACGTCGCAACAGCGATCGGTGCTA 1091
Qy      241 LeuPheArgArgIleAlaAlaSerThrIleAlaPheIleSerTyrLeuGlySerPheAla 260
Db      1092 CTGTTCAAGAGGATCGCAGCGAGCACCATCGCCTTCATTTCTATTTAGGCGAGCTTTGCA 1151
Qy      261 AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 280
Db      1152 GCGCAACTGAAANATAGCAGCAGCAGCAGTAGCAGCAGCAACAGCAACACACGACGAGC 1211
Qy      281 ThrGlnIleLeuAsnGlyLeuAsnLysHisSerTrpIlePheLeuLeuIleTyrLeuAsn 300
Db      1212 ACGCAATATTAACCGGACTTAATAACACTCATGGATATTTTATTGATATATTGTAAT 1271
Qy      301 LeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArgLeuLeuHisAspLeuLeu 320
Db      1272 TTATCTGCTAAAGTTTGCCTAGCAGGATATCATGAAAGAGACTGTTACAGATCTTTTG 1331
Qy      321 AspProTyrAsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSer 340
Db      1332 GATCCTTATAATACACTAGAACGTCCTGCTCAATGAATCGGACCCGTTACAAATTAAGC 1391
Qy      341 PheGlyLeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuLeuValThr 360
Db      1392 TTTGGTTTAACTTTAATGCAAAATATCGATGTCGACGAGAAAAATCAATTCTGACTACT 1451
Qy      361 AsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyr 380
Db      1452 AATGTGGTTTAAACCTGGAGTGAACGACATGAATCTCCGCTGGAACACCTCCGACAT 1511
Qy      381 GlyGlyValLysAspLeuArgIleProProHisArgIleTrpLysProAspValLeuMet 400
Db      1512 GCGGAGTTAAGGATCTGCGAATACCCGCGCATCGCATCTGGAAGCGGACGTCGTGATG 1571
Qy      401 TyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgAsn 420
Db      1572 TACAACAGTTCGGATGAGGATTTGACGGCACTTACCAGACGAACTGCTGGTGGCGAAC 1631
Qy      421 AsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIle 440
Db      1632 AACGGCTGCTGTATAGTTCCCGCGGGATCTTCAAGTCGACGTGCAAGATCGACATC 1691
Qy      441 ThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAsp 460
Db      1692 ACGTGGTTCCCTTCGATGACCGCGGTGCGAGATGAAGTTCCGCGAGTTGGACCTACGAC 1751
Qy      461 GlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGlyAspIleSerSerTyrVal 480
Db      1752 GGATTCCAGCTGGATTTACAATTAAGATGAATGCGGTGATATCAGGAGTTTACGTTG 1811
Qy      481 LeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsn 500
Db      1812 CTCACGCGGAGTGGGAACCTACTGGGTGTGCCCGCAACCGTAACGAGATCTATTACAC 1871
Qy      501 CysCysProGluProTyrIleAspIleThrPheAlaIleIleIleArgArgThrLeu 520

```

```
Db 1872 TGCTGCCGGAGCCTATATAGACATACCTTCGCCATCATCCGCGGAGCAACTG 1931
Qy 521 TyrTyrPhePheAsnLeullelleProCysValLeullelleAlaSerMetAlaLeuLeuGly 540
Db 1932 TACTATTTCCTCAACCTGATCATACCTTGCTAGTATGCTCCATGGCCTTGCTCGG 1991
Qy 541 PheThrLeuProProAspSerGlyGluLeuSerLeuGlyValThrLeuLeuSer 560
Db 1992 TTACCTCTGCCCGCAGATTGGGTGAAATAATTCGCTGGGTGTATCCATCTTGCTCTCG 2051
Qy 561 LeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValPro 580
Db 2052 CTGACCGGTGTTCTGAATAGTTGGCAGACATCGCGCTACTTCGATGGCGTGCCA 2111
Qy 581 LeuTrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMetSerArgPro 600
Db 2112 TTGTGGATAGCATCGTGTGTTTTGTGCTGGCTGCATGATATTCGAATGAGTCGCCCA 2171
Qy 601 GlyArgProLeuLeuLeuGluPheProThrThrProCysSerAspThrSerSerGluArg 620
Db 2172 GGACGACCGCTGATCTTAGAGTTCGCCACACGCCCTGTTCGGACATCTCCCGACGG 2231
Qy 621 LysHisGlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSerLeuLeuAla 640
Db 2232 AAGCACCAGATACTCTCCGAGTTGAGCTGAAGAGCGCTCGTGAATCGCTGCTGCC 2291
Qy 641 AsnValLeuAspIleAspAspAspPheArgHisAsnCysArgProMetThrProGlyGly 660
Db 2292 AACGTACTAGACATCGATGATGACTTCGGGCACAAATTTGTCGCCCATGAGCCCGCGGA 2351
Qy 661 ThrLeuProHisAsnProAlaPheTyrArgThrValTyrGlyGlnGlyAspAspGlySer 680
Db 2352 AACTGCCACACACACCGCGCTTCTATCGACGGTTTTATCGACAAAGCGACGATGGCAGC 2411
Qy 681 IleGlyProIleGlySerThrArgMetProAspAlaValThrHisHisThrCysIleLys 700
Db 2412 ATTGGGCCAATTGGCAGCACCAGATCCGAGTCCGCTCACCACATCATAGTCATCAAA 2471
Qy 701 SerSerThrGluTyrGluLeuGlyLeuLeuLysGluLeuArgPheIleThrAspGln 720
Db 2472 TCATCAACTGAATATGAATAGGTTTAACTTAAAGGAAATTCGTTTATAACTGATCAG 2531
Qy 721 LeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetVal 740
Db 2532 CTACGTAAAGATGACGAGTGCATGACATTCGCAATGATTTGGAAATTTGACGCTATGGTC 2591
Qy 741 ValAspArgLeuCysLeullellePheThrMetPheAlaIleLeuAlaThrIleAlaVal 760
Db 2592 GTTGACAGACTGTGCCTTATCATATTCACAAATGTTGCATATATTAGCCACATAGCTGTA 2651
Qy 761 LeuLeuSerAlaProHisIlelleValSer 770
Db 2652 CTACTATCGGCACCATATATTATTGCTCG 2681
```

```
RESULT 2
AX009610 LOCUS
DEFINITION Sequence 1 from Patent EP0962528.
ACCESSION AX009610
VERSION AX009610.1 GI:9996842
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1
REFERENCE Adamczewski,M.D., Schulte,T.D. and Oellers,N.D.
AUTHORS Nucleic acids encoding acetylcholin-receptor subunits from insects
TITLE Patent: EP 0962528-A 1 08-DEC-1999;
JOURNAL BAYER AG (DE)
FEATURES Location/Qualifiers
```

source

```
1. 2886
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"
372..2684
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC07499.1"
/db_xref="GI:9996843"
/db_xref="RENTREMBL:CAC07499"
/translation="MKNAQLKLEVDDELDLAVLRLAHCSNFSSTSTRTSSNOEH
NQQLTLPESLSLTKHSNITASECHNSQQQEPASKQEDVANHGSRNQDQHLQQLDSS
NMSPKTAATAAGDEATTQQTNTLRCARKQRRLRRRRKRPATNETDIDKQQQL
SMSPFKTKSTDTYSTPATTSCTATYMQCRASDNEFSIPISRHDRVSTATFAWVL
VLQGLVLSQQWQLHQVRSVLLFRRIAASIAFI SYLGSFAAQKLNSSSSSSNSS
NNSSTQTLNGLNKHSMIFILIYLNLSAKVCLAGYHEKRLHLDLDPYNTLERPVNES
DPLQGLFGLTMOI IDVDEKNQLLVNWLKLEWDMNLRWNTSDYGVKDLRTPPHR
IKPDDVLMYNSADGPDGTQTNVVRNNGSLYVPGI FKSTCKIDITWPFEDQRC
EMKPGSWYDGFQDLQDQETGDISSYVNGEWELLGVFGRNRIYINCCPEPIYD
ITFAIIRRLTYFFNLIIPCVLIASMALLFTLPDPSGEKLSGLVTILLSLTFLVN
MVAETPATSDAVPLMIRIVLCWLPMILRMSRGRPLILEFTTTCSDTSSBRKHQI
LSDVELKRSKSLANVLDDDFRHNCRPMTGGTLPHNPAFYRTVYQGGDGSIG
PIGSTMPDAVTHHTCIKSTSEYELGLILKEIRFITDQLRKDDDECNDIANDWKFAMV
VDRCLLIIFTFMFAILATIAVLLSAPHIVS"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 7,886-286 Length: 2886
Score: 4043.00 Matches: 770
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-303-232-2 (1-770) x AX009610 (1-2886)

Qy 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
Db 372 ATGAAAAATCACAACACTGAACTGACTGAAGTTGACGATGATGAGCTGTGGTCAGTA 431
Qy 21 ArgLeuAlaHisCysSerSerAsnPheSerSerSerSerSerSerThrArgThrSerSer 40
Db 432 AGATTAGCGCACTGCGACGACAACTTTAGCAGCAGTAGCAGCACAGAACCCAGCAGC 491
Qy 41 AsnGluArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
Db 492 AACGAGAGGCACACACGACAACTCACAACACTGCAACCAAGGAGCTTAAGTACAAAACAC 551
Qy 61 HisSerAsnIleAlaSerGluGlnHisAsnSerGlnGlnGlnGlnProAlaSerLysAsp 80
Db 552 CACAGCAACATTGTCAGCGAGCAGCACATAGCCAGCAACAGGAGCCAGCATCGAAGGAC 611
Qy 81 GluAspValAlaAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp 100
Db 612 GAGGATGTAGCAACCAACCGGTAGAACCAATGACACGACGACGATCTGCAACAGTAGAC 671
Qy 101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaAlaAlaAlaGlyAspGluAla 120
Db 672 AGCAGCAACATGTTGTGCGCAAGACAGCGCAGCAACCTGCTGCCGCGCATGAAGCA 731
Qy 121 ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
Db 732 ACAACCCAAACCAACCAACCAATAGACTGTGTGTCAGCAAGCAGCAACGATTCGTCGCG 791
Qy 141 ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLysLysGlnGlnGlnLeu 160
Db 792 CGACGAAAAGAAAACACAGCAACCCCAACAGAAACAGATATCAAGAAACAAACAGCACTT 851
Qy 161 SerMetProProPheLysThrArgLysSerThrAspThrTyrSerThrProAlaAlaThr 180
Db 852 AGCATGCCCTCCCTTCAAAACGGCGAAATCCACGACACCTACAGCACACACGACGACACA 911
Qy 181 ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGluPheSerIle 200
```


FEATURES	Location/Qualifiers		Score:	4014.50	Matches:	770
	1..2907		Percent Similarity:	95.42%	Conservative:	0
Source	/organism="Drosophila melanogaster"		Best Local Similarity:	95.42%	Mismatches:	0
	/mol_type="mRNA"		Query Match:	99.30%	Indels:	37
	/db_xref="taxon:7227"		DB:	3	Gaps:	1
	/chromosome="2"					
gene	/map="2L; 34B4-5"		US-09-303-232-2 (1-770) x AF272778 (1-2907)			
	/dev stage="embryo"					
	1..2907					
	/gene="nAcRalpha-34E"					
CDS	/allele="A"		Qy	1	MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal	20
	289..2712		Db	289	ATGAAAAATGCACAACTGAAACTGACTGAAGTTGAACGATGATGAGCTGTGGCTGGCAGTA	348
	/gene="nAcRalpha-34E"		Qy	21	ArgLeuAlaHisCysSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer	40
	/note="ion channel; neurotransmitter transmembrane receptor"		Db	349	AGATTAGCGCACTGCAGCACTTTAGCAGCAGTAGCAGCACAAGAACACACGAGCAGC	408
	/codon_start=1		Qy	41	AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis	60
	/product="nicotinic acetylcholine receptor Dalpha5 subunit"		Db	409	AACCAGAGGCACACACCACTCAACAACTGCAACCAAGGAGCTTAACTACAAAACAC	468
	/protein_id="AAM13390.1"		Qy	61	HisSerAsnIleAlaSerGluGlnHisAsnSerGlnGlnGlnGlnProAlaSerLysAsp	80
	/db_xref="GI:20152840"		Db	469	CACAGCAACATTGCAAGCAGCAGCACAATAGCCAGCAACAGAGCCAGCATCGAAGGAC	528
	/translation="MKNAQLKLTVEVDDDELAVRLAHCSNFSSTSSSTRTTSSNORH NQQLTTLQPRSLSTKHSNIASEQHNSSQQQEPASQEDVANHGKSNDDQQLHQQLDSS NMLSPKTAARATAGDEATTQQPTNIRKLCARKQRRRRKRPKPTNETDIKKQQQL SMPEFKRSTDTYTPAATTSCTPTATMQCRASDNEFSPISRHRVNSTATFAWVLH VLQVLLSQOQLHVQQRSLVFRRIAATIAFISYLGSAALQDLNLSSTSSSSSSSS NNSSTQTLNGLNKHSTFLLIYLNLSAKVCLAGYHEKRLHLDDLPYNTLERPVLNES DPIQLSGFLTLMQI IDVEKNQLLVNWLKLEWDMNLFWNTSDYGGVKDLRIPEHR IMKPDVLYNSADEGPDGTQTVNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRC EMKFGSWTDGFDLQLODETGDISSYVINGEWELLGVPPGRNEIYINCCEPEYID ITPAIIRRTLLYFNLLIPCVLIASMLLFTLPDSEKLSLGVTTILSLTVFIN MVAETPATSADVPLGTFCNFMFVASSVSTILILNYHHRNADTHSEMSEWIRIVF LCWPLWLRSPGRLLIEFTPTPCSDTSSEKHOILSDVELKRSKSLLANVLDI DDDFRNCMPTEGGTLPHNPAFVRTVYGGDDGSGICPIGSTTMPDPAVHTTCKIST EYELGILLKEIRITQLRKDDCEINDWKFAAMVVDRLCLITFMFAILLIIVL LSAPHIVS"		Qy	101	SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla	120
	373		Db	589	AGCAGCAACATGTTGCGCCAAAGACAGCCGCGAGCACTGCTGCCGCGCATGAAGCA	648
variation	/gene="nAcRalpha-34E"		Qy	121	ThrThrGlnGlnProThrAsnIleArgLeuLysCysAlaArgLysArgGlnArgLeuArgArg	140
	/note="compared to B allele"		Db	649	ACAACCCCAACACAAACAAATAAGACTGTGTGCACGCAAGCAACAACGATTGGCTGC	708
variation	/replace="a"		Qy	141	ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLysLysGlnGlnGlnLeu	160
	375		Db	709	CGACGAAAGAAACACGACCCCAACGAAACAGATATCAAGAAACACGCAACTT	768
variation	/gene="nAcRalpha-34E"		Qy	161	SerMetProProPheLysThrArgLysSerThrAspThrThrThrProAlaAlaThr	180
	/note="results in phenylalanine to isoleucine substitution; compared to B allele"		Db	769	AGCATGCTCCCTTCAAAACGCGCAATCCACGGAACCTTACAGCACACACGAGCAGAC	828
variation	827		Qy	181	ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGluPheSerIle	200
	/note="results in threonine to isoleucine substitution; compared to B allele"		Db	829	ACCAGTGTCCGACAGCCACCTACATGCAATGTGAGCCAGCCAGCAATGAGTTT	888
variation	/replace="t"		Qy	201	ProIleSerArgHisAspArgValSerThrAlaThrPheAlaTrpValLeuHisValLeu	220
	1079..1080		Db	889	CCGATATCGAGACATGATAGATATCCCGGCCACATTCGCTGGGTGTTCATGTGCTG	948
variation	/gene="nAcRalpha-34E"		Qy	221	GlnValLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu	240
	/note="results in lysine to arginine substitution; compared to B allele"		Db	949	CAGTGCTGCTCGTGTGCGTGCACAGTGGCACTTCAGTGCACACGAGTCGGTGCTA	1008
variation	1081..1086		Qy	241	LeuPheArgArgIleAlaAlaSerThrIleAlaPheIleSerTyrLeuGlySerPheAla	260
	/note="results in deletion of asparagine and serine; compared to B allele"		Db	1009	CTGTTTCAAGAGGATCGCAGCAGCACCATCGCTTCATTTCCTATTATTAGGACGCTT	1068
variation	/replace="g"		Qy	261	AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer	280
	1119..1120		Db	1069	GCGCAACTGAAATAAGTAGCAGCAGCAGTAGCAGCAGCAACAGCAGCAACACGAGC	1128
variation	/gene="nAcRalpha-34E"		Qy	281	ThrGlnIleLeuAsnGlyLeuAsnLysHisSerTrpIlePheLeuLeuIleTyrLeuAsn	300
	/note="results in serine, glycine, and asparagine insertion; compared to B allele"		Db	1129	ACGCAATATTAAACGACTTAATAAACACTCATCGATATTTTATTGATATATTTGAAT	1188
variation	1121		Qy	301	LeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArgLeuLeuHisAspLeuLeu	320
	/note="results in asparagine to serine substitution; compared to B allele"		Db	1189	TTATCTGCTAAAGTTTTCCTAGCAGGATATCATGAAAGAGAGACTGTTACACGATCT	1248
variation	/replace="g"		Qy	321	AspProTyrAsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSer	340
	9..64E-284					
ORIGIN						
Alignment Scores:						
Pred. No.:						

```

1249 GATCCCTTATATACACTAGAACGTCCTCCGTTCTCAATGAATCGGACCGCTTACAATTAAGC 1308
341 PheGlyLeuThrLeuMetGlnIleAaspValaspGluLysAsnGlnLeuLeuValThr 360
1309 TTTGGTTAACTTTAATGCAAAATATCGATGTGACGAGAAAAATCAATTTGCTAGTCACT 1368
361 AsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyr 380
1369 AATGTGTGGTTAAACCTGGAGTGAACGACATGAATCTCCGCTGGAAACACCTCCGACTAT 1428
381 GlyGlyValLysAspLeuArgIleProProHisArgIleTrpLysProAspValLeuMet 400
1429 GCGCGAGTTAAGGATCTCGCAATACCGCGCATCGCATCTGGAAGCGGACGCTCTGATG 1488
401 TyrAsnSerAlaaspGluGluPheAaspGlyThrTyrGlnThrAsnValValValArgAsn 420
1489 TACAACAGTGGGATGAGGATTTTGACGGCACCCTACCAGAGCAAGCTGGTGGTGGGAAC 1548
421 AsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIle 440
1549 AACGGCTCGTGTCTATACGTTCCCGCGGGGATCTTCAAGTCGACGTCGACAGATCGACATC 1608
441 ThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAsp 460
1609 ACGTGGTTCCCTTCGATGACACCGCGTGGAGATGAAGTTCCGCGAGTTGGACCTACGAC 1668
461 GlyPheGlnLeuAspLeuGlnLeuAspGluThrGlyGlyAspIleSerSerTyrVal 480
1669 GGATTCGACGTGGATTTACAATTAAGATGAATCGCGGTGATATCAGCAGGTACGTTAGCTG 1728
481 LeuAsnGlyGluTrpGluLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsn 500
1729 CTCAACGGCGAGTGGGAACCTACTGGGTGTGCGCGCAACGTAACGAGATCTATTACAC 1788
501 CysCysProGluProTyrIleAaspIleThrPheAlaIleIleAaspArgArgThrIleu 520
1789 TGCTGCCGGAACCTATATAGACATCACCTTCGCCATCATCTCCGCGGACGAACTG 1848
521 TyrTyrPhePheAsnLeuIleIleProCysValLeuIleAlaSerMetAlaLeuLeuGly 540
1849 TACTATTCTTCAACCTGATCATACCTTGTGTACTGTGATTCCTCATGCGCTTCTCGGA 1908
541 PheThrLeuProProAspSerGlyLysLeuSerLeuGlyValThrIleLeuLeuSer 560
1909 TTCACCTCGCGCAGATTCGGGTGAAAAATATCGCTGGGTGTTACCATCTTCTCTCG 1968
561 LeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValPro 580
1969 CTGACCGGTGTTCTGAATATGTTGCCGAGACAATGCCGGCTACTTCCGATCGCGTGCCA 2028
581 Leu----- 581
2029 TTGCTGGGTACATATTTCAATTGCATATGTTATGTGTAGCTTCATCCGTTGTGTCAACG 2088
582 ----- 582
2089 ATTTTAAATTAATATCATCATCGAATGCTGATACGACGCAAAATGTCGGAATGGATA 2148
584 ArgIleValPheLeuCysTyrTrpLeuProTrpIleLeuArgMetSerArgProGlyArgPro 603
2149 CGCATCGTGTGTTTGTGCTGCTGCGCATGGATATTGCGAATGAGTCGCGCCAGGACGCGC 2208
604 LeuLeuLeuGluPheProThrThrProCysSerAspThrSerSerGluArgLysHisGln 623
2209 CTGATCTTAGATTCCCGACACGCGCTTTCGGACACATCTTCGAGCGGAAGCACCAG 2268
624 IleLeuSerAspValGluLeuLysGluArgSerLysSerLeuLeuAlaAsnValLeu 643
2269 ATACTCTCCGACGTTGAGCTGAAGAGCGCTCGTCAAAATCGCTGCTGCCCAACGTAATA 2328
644 AspIleAspAspPheArgHisAsnCysArgProMetThrProGlyGlyThrLeuPro 663
```

```

2329 GACATCGATGATGACTTCGGGCACAAATTGTGCCCCCATGACGCCCGCGGAACACTGCCA 2388
664 HisAsnProAlaPheTyrArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyPro 683
2389 CACAACCGCGCTTTTCATCCACGTTTATGGACAAGCGCAGATGCGCAGCATTTGGGCCA 2448
684 IleGlySerThrArgMetProAspAlaValThrHisThrCysIleLysSerSerThr 703
2449 ATTGGCAGCACCGAATGCGGATGCGGTCAACCATCATACGTGTCATCAATCATCAACT 2508
704 GluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLys 723
2509 GAATATGAATTAGCTTTAATCTTAAAGCAAAATTCGCTTTATACTGATCAGTACGTAA 2568
724 AspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArg 743
2569 GATGACGAGTGCATGACATTCGCAATGATTCGAAATTTGAGCATATGTTGTTGACAGA 2628
744 LeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSer 763
2629 CTGTGCTTATCATATTCAATGTTCCGAATGTTAGCCCAATAGCTGACTACTATCA 2688
764 AlaProHisIleIleValSer 770
2689 GCACCATATATTGTTCTCG 2709
```

RESULT 4
AY036613 2834 bp mRNA linear INV 30-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalpha5
DEFINITION subunit (nAcRalpha34E) mRNA, nAcRalpha34E-B allele, complete cds,
alternatively spliced.
ACCESSION AY036613
VERSION AY036613.1 GI:20340268
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2834)
AUTHORS Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
TITLE Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
JOURNAL Genetics 160 (4), 1519-1533 (2002)
MEDLINE 21969411
PUBMED 11973307
REFERENCE 2 (bases 1 to 2834)
AUTHORS Grauso, M. and Sattelle, D.B.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-2001) MRC-FGU Human Anatomy and Genetics,
University of Oxford, South Parks Road, Oxford OX1 3QX, UK

FEATURES
Location/Qualifiers
1..2834
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="34E4-34E5"
/dev_stage="embryo"
1..2834
/gene="nAcRalpha34E"
/allele="B"
289..1737
/gene="nAcRalpha34E"
/note="ion channel; neurotransmitter transmembrane
receptor; exon 5 is excluded due to exon skipping, which
generates a loss of reading frame and a truncated
polypeptide; alternatively spliced"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha5"

gene
CDS

subunit"
 /protein id="AAK67256.1"
 /gb_xref="GI:20340263"
 /translation="MKNAQLKLTVEVDDDELWLAFLHLCSSNISSSSSTRRTTSNQHRH
 NQQLITLQPSLSTKSHNSIASEQHNSQQQEPASKDQVANHGRSNDQOHLQOLDSS
 NMLSPKTAATAAAGDEATTOOPTNIRLCARQRRLRRRRKRPATNETDIKKQOOL
 SMPFKRKSTDTYSTFAAITSCPTATYMCRASDNFSPISRHRVSTATFAWLH
 VLOVLLSQOQLHVOORSVLLFRRIAASITAFISYLGSPAOLRVSSSSSSSNS
 GNSSTOILGNLKHSMFLILYLNLAKVCLAGVHEKRLHLHLLDPYNTLERVPLNE
 SDPLQLSGLTLMOIIDVDEKNQLLVTNWLKLEWMDNLRKNTSDTGGVADLRIPPH
 RIWFDVLMYNSADEGFGTYQTNVVRNNGSLYVPPGIFPKSTCKIDITWFFPDDOR
 CEMRFGSMTYDGFQVCPANVTRSIITAARNP"

ORIGIN

Alignment Scores:

Pred. No.: 2,09e-268 Length: 2834
 Score: 3804.50 Matches: 738
 Percent Similarity: 91.71% Conservative: 3
 Best Local Similarity: 91.34% Mismatches: 4
 Query Match: 94.10% Indels: 64
 DB: 3 Gaps: 3

US-09-303-232-2 (1-770) x AY036613 (1-2834)

Qy 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
 Db 289 ATGAAAAATGCACAACTGAAACTGACTGAACTTCACGATGATGAGCTGTGGCTGGCAGTA 348
 Qy 21 ArgLeuAlaHisCysSerSerAspPheSerSerSerSerThrArgThrSerSer 40
 Db 349 AGATTAGCGCACTCCAGCAGCAACATCAGCAGCAGTAGTAGCAGCACAAAGAACCCAGCAGC 408
 Qy 41 AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
 Db 409 AACAGAGGCACACACAGCAACTCACCAACTGCACCAAGGAGCTTAAGTACAAAAAC 468
 Qy 61 HisSerAsnLeuAlaSerGluGlnHisAsnSerGlnGlnGlnGluProAlaSerLysAsp 80
 Db 469 CACAGCAACATTCAGAGCAGCAGCACAATAGCCAGCAACAGAGCCAGCATCGAAGGAC 528
 Qy 81 GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp 100
 Db 529 GAGGATGTAGCCAAACCAACCGTAGAAGCAATGACAGCAGACGGCATCTGCAACAGCTAGAC 588
 Qy 101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla 120
 Db 589 AGCAGCAACATGTTGTCCCAAGACAGCCGCGAGCAACTGTGCGCGGAGTAGAGCA 648
 Qy 121 ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
 Db 649 ACAACCCCAACCAACAAACATAAGACTGTGTGACGCGCAGCAGCAACGATTGCGTGC 708
 Qy 141 ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLysLysGlnGlnLeu 160
 Db 709 CGACGAAAGAAAGAACCGAACCCCAACCGAACACAGATATCAAGAAACCAACAGCAACT 768
 Qy 161 SerMetProProPheLysThrArgLysSerThrAspThrThrSerThrProAlaAlaThr 180
 Db 769 AGCATGCCCTCCCTTCAAAAGCGCGCAATCCACGACACCTACAGCAGCAGCAGCAATA 828
 Qy 181 ThrSerCysProThrAlaThrTyMetGlnCysArgAlaSerAspAsnGluPheSerIle 200
 Db 829 ACCAGCTGTCGACAGCACCTACATGCAATGTCGAGCCAGCAGCAATGAGTTCAGTATT 888
 Qy 201 ProIleSerArgHisAspArgValSerThrAlaThrPheAlaTrpValLeuHisValLeu 220
 Db 889 CCAGATCCAGACATGATAGTATCCAGCCACCATTCGCTGGGTGTGATGTGCTG 948
 Qy 221 GlnValLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu 240
 Db 949 CAGGTGCTGCTGTGCTGCAACAGTAGTGCACACTTCACGTGCAACAGCGATCGGTGCTA 1008
 Qy 241 LeuPheArgArgIleAlaAlaSerThrIleAlaPheIleSerThrLeuGlySerPheAla 260

Db 1009 CTGTTTCAAGAGATCGCAGCAGCACCATCCCTTCAATTCCTATTTAGGCAGCTTTGCA 1068
 Qy 261 AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 279
 Db 1069 GCGCAACTGAGGAGCAGCAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1128
 Qy 280 SerThrGlnIleLeuAsnGlyLeuAsnLysHisSerTrpIlePheLeuLeuIleTrpLeu 299
 Db 1129 AGCAGCGAAATATTAACCGGCTTAATAAACACCTCATGATATTTTATTGATATATTG 1188
 Qy 300 AsnLeuSerAlaLysValCysLeuAlaGlyTyHisGluLysArgLeuLeuHisAspLeu 319
 Db 1189 AATTATCTGCTAAAGTTTGCCTAGCAGGATATCATGAAAAGAGACTGTTTACAGATCTT 1248
 Qy 320 LeuAspProTyAsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeu 339
 Db 1249 TTGGATCTTATATATACACTAGAACGTCGCTTCTCAATGATCGGACCCGTTTACAATTA 1308
 Qy 340 SerPheGlyLeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuVal 359
 Db 1309 AGCTTTGGTTTAACTTTAATGCAATATATCATGTGACGAGAAAAATCAATTGCTAGTC 1368
 Qy 360 ThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAsp 379
 Db 1369 ACTAATGTGTGTTTAAACCTGGAGTGGAGACGACATGAATCTCCGCTGGAACACCTCGAC 1428
 Qy 380 TyrGlyGlyValLysAspLeuArgIleProProHisArgIleTrpLysProAspValLeu 399
 Db 1429 TATGGCGAGTTAAGGATCTGCCAATACCCCGCATCGCATCTGGAAGCCGAGCGTCTG 1488
 Qy 400 MetTyrsnSerAlaAspGluGlyPheAspGlyThrTyTrpGlnThrAsnValValArg 419
 Db 1489 ATGTACAAACAGTCGGATGAGGATTTGACGGCACCTTACCAGACGAACGTGTGTGGCG 1548
 Qy 420 AsnAsnGlySerCysLeuTyTrpValProProGlyIlePheLysSerThrCysLysIleAsp 439
 Db 1549 AACACCGCTGTGTCTATACGTTCCGCGGGGATCTTCAAGTCGACGTCGCAAGATCGAC 1608
 Qy 440 IleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerThrTyTrp 459
 Db 1609 ATCAGCTGGTTCCTTCCTTCGATGACCGAGGTGGAGATGAAGTTCGCGAGTTGACCTAC 1668
 Qy 460 AspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGlyAspIleSerSerTy 479
 Db 1669 GACGGATTCCA----- 1679
 Qy 480 ValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyTrp 499
 Db 1680 -----GGTGTGCGCGGCAACGTAACGAGATCTATTAC 1712
 Qy 500 AsnCysCysProGluProTyTrpIleAspIleThrPheAlaIleIleIleArgArgArgThr 519
 Db 1713 AACTGTCGCGGAAACCTATATAGACATCACTTCGCCATCATCATCCGCCAGCAGACA 1772
 Qy 520 LeuTyTrpPhePheAsnLeuIleIleProCysValLeuIleAlaSerMetAlaLeuLeu 539
 Db 1773 CTGTACTATTTCTTCAACCTGATACCTTGTGTACTGATTCGCTCCATGGCCTGTGCTC 1832
 Qy 540 GlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeu 559
 Db 1833 GGAATCACCTGCGCGCAGATTGCGGTGAAAAATATATCGCTGGGTGTTTACCATCTTGCTC 1892
 Qy 560 SerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaVal 579
 Db 1893 TCGCTGACCGTGTTCGTGATATGTTGTCGAGACAAATCCCGCTACTTCCGATGCGGTG 1952
 Qy 580 ProLeu----- 581
 Db 1953 CCATTGCTGGGTACATATTTCAATTGCAATAATGTTTATGGTAGTTCATCCCGTGTGTGCTA 2012
 Qy 582 -----Trp 582

```
Db      2013  ACGATTATTAATTAATTCATCATCGAAATGCTGATACGACGAAATGTCGCAATGG 2072
Qy      583  ILAArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMetSerArgProGlyArg 602
Db      2073  ATACGCATCGTGTGTTGCTGCTGCCATGATATTCGAAATGAGTCGCCACGACGA 2132
Qy      603  ProLeuIleLeuGluPheProThrThrProCysSerAspThrSerSerGluArgLysHis 622
Db      2133  CCGCTGATCCTAGAGTTCGACACACCGCCCTGTCGACACATCTCGACGCGAAGAC 2192
Qy      623  GlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 642
Db      2193  CAGATACTCCGAGGTTGAGCTGAAAGAGCGCTCGTCGAAATCGCTGCGCCACGTA 2252
Qy      643  LeuAspIleAspAspPheArgHisAsnCysArgProMetThrProGlyGlyThrLeu 662
Db      2253  CTAGACATCGATGATGACTTCGCGCAAAATTTGCGCCCATGACGCCCGCGGACACATG 2312
Qy      663  ProHisAsnProAlaPheThrArgThrValTyrGlyGlnGlyAspAspGlySerIleGly 682
Db      2313  CCACACAAACCGCGCTTCTATCGACAGGTTTATGACACAGCGAGATGGCAGCATGGG 2372
Qy      683  ProIleGlySerThrArgMetProAspAlaValThrHisHisThrCysIleLysSerSer 702
Db      2373  CCAATTGGCAGCACCCGAAATGCGGATGCGGTCAACCATCATACGTGCATCAAAATCATCA 2432
Qy      703  ThrGluTyrGluLeuGlyLeuIleLeuLysGluLeuArgPheIleThrAspGlnLeuArg 722
Db      2433  ACTGAATATGAATAGGTTTAAATCTTAAAGAGAAATTCGCTTTTAACTGATGACGTACGT 2492
Qy      723  LysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAsp 742
Db      2493  AAAGATGACGAGTGCATGATGATGCGCATGATGCGCAATGATGGAATTTGAGCTATGCTGTAC 2552
Qy      743  ArgLeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeu 762
Db      2553  AGACTGCTCCTATCATATTACAAATGTTGCAATATTCGCAATATTAGCCAAATAGCTGTACTACTA 2612
Qy      763  SerAlaProHisIleIleValSer 770
Db      2613  TCAGCACACATATTATTGTCTCG 2636

RESULT 5
DME554210 1683 bp mRNA linear INV 01-APR-2003
LOCUS      Drosophila melanogaster mRNA for nicotinic acetylcholine receptor
DEFINITION subunit Dalpha7 (nAcRalpha-18C gene).
ACCESSION  AJ554210
VERSION     AJ554210.1 GI:29466436
KEYWORDS    nAcRalpha-18C gene; nicotinic acetylcholine receptor subunit
SOURCE      Dalpha7.
ORGANISM    Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1683)
MILLAR,N.S.
Direct Submission
Submitted (31-MAR-2003) Millar N.S., Department of Pharmacology,
University College London, Gower Street, London, WC1E 6BT, UNITED
KINGDOM
FEATURES
source      Location/Qualifiers
1..1683
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1..1683
/gene="nAcRalpha-18C"
26..1654
/gene="nAcRalpha-18C"
/codon_start=1
/product="nicotinic acetylcholine receptor subunit
```

```

Dalpha7"
/protein_id="CAD86936.1"
/db_xref="GI:29466437"
/translation="MSFPDPHSLPEATANGRMIVYGLGLIMIPACAAGPHEKRLIH
ALLDNNLSERPVNNSDPQLSFGILTMQIIDVDEKNOLLITNWLKLEWNNMLSW
NSSEFGRLRIPPHRLWKPDMYNSADEGFDGTATNVVVRNNGSCLYVPPGIFK
STKIDITWFFDQRCMKFGSYDGLDQLQDEAGDISSTITNGEWDLGVP
GKENIYNCCPEPYDITPAILIRKTLVFFNLIVPCVLIASMLLGLTLPDSGE
KLSGLTLLSLVFLNVAETWPAATSDAVLLGKFCNCFMFWASSVSTILVLNTH
HRNPDTHEMSEIRVILFWLPCILRMORPGQVCEPPPPSSSSASGEKKQIQN
VEUKERSKSLLANVLIDDDDFRCNRCASATLPHQTYRTMYRQGDGSGVPGPA
GPVDRGLHEAHSHTLSAEYELALILKELWITEQLKKEDEBTSITDRDKWFAAV
VDRGLGIITFTIATLAVLFSAPHFIVSGVRG"
26..131
sig_peptide
/gene="nAcRalpha-18C"
132..1651
mat_peptide
/gene="nAcRalpha-18C"
/product="nicotinic acetylcholine receptor subunit
Dalpha7"

Alignment Scores:
Pred. No.:      3.85e-136      Length:      1683
Score:          1991.00      Matches:      382
Percent Similarity: 80.95%      Conservative: 43
Best Local Similarity: 72.76%      Mismatches: 46
Query Match:     49.25%      Indels:      54
DB:              3          Gaps:      8

US-09-303-232-2 (1-770) x DME554210 (1-1683)
Qy      295  LeuLeuIleTyr-----LeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHis 311
Db      80  ATGCTGGTCTATGCGCTGGGACTTTTAAATTATGATACCGGCTTGTGCGCTGGACCCCAT 139
Qy      312  GluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeu 331
Db      140  GAGAAGCGGCTACTCCAGCCCTTCTGACAACTACACAGCCGCGGAGCGTCCGGTGGTC 199
Qy      332  AsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIleAspVal 351
Db      200  AATGAATCCGATCCATTGCCAATCGAGCTTCGAGCTTACACTCATCTGCAGATTATCGATGTG 259
Qy      352  AspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMet 371
Db      260  GACGAAAGAAATCAACTGCTTTATACGAATATTTGGCTCAAAATGGAATGGAACGATATG 319
Qy      372  AsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProHis 391
Db      320  AATCTTCGATGGAATTCGAGTGAGTTCGGTGGTTCGGGATCTCGAATTCGCCACAT 379
Qy      392  ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db      390  CGCTATGGAAACCGATGCTACTGATGTACACAGTGGCGAGGGCTTCGATGGAACG 439
Qy      412  TyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIle 431
Db      440  TAGCCCAAAATGTGGTGGTTCGCAATTAATGGAGCTGTCTGTACGTACCGCAGGTATA 499
Qy      432  PheLysSerThrCysLysIleAspIleThrPheProPheAspAspGlnArgCysGlu 451
Db      500  TTTAAGTCAACGTGTAAAGATGACATTACGTTGTTTCCATTCGAGATCGAGATGTGAA 559
Qy      452  MetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGlu 471
Db      560  ATGAAATTTGGTTCTGTGACCTACGATGGGTTTCAGTTGGACCTGCTGTCAGGACGAA 619
Qy      472  ThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValPro 491
Db      620  GCTGTGGCGACATTTCTAGCTTTTAAACCAATGCGAATGGGATGTTGTAGTGTGCC 679
Qy      492  GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
```



```

Db 680 GGTAACGAATGAATCTACTATAATGCTGCCAGAACCTTATTATGACATAACATTC 739
QY 512 AlallelleleAqArgThrLeuTyrPhePheAsnLeullelleProCysVal 531
Db 740 GCCATTTGATAGCGGCAACAGTGTACTATTTTCAATCTCATTTGTCGCGTGA 799
QY 532 LeulleAlaSerMetAlaLeuLeuGlyPheThrLeuProAspSerGlyGlyLeu 551
Db 800 CTGATCGCCTCATGGCATGCTAGGGTTTACACTGCCACAGATCTGGTGAAGCTT 859
QY 552 SerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValalaGluThr 571
Db 860 TCGCTTGAGGTACAAATCTATATCGCTTACAGTCTTCTCAACATGGTGGCGGAACA 919
QY 572 MetProAlaThrSerAspAlaValProLeu----- 581
Db 920 ATGCGCGGCACTCGATGCGGTACCGCTGCTCGGTAAGTATTTCAATTGCATTATGTTT 979
QY 581 ----- 581
Db 980 ATGCTGGCTCATCAGTTGTGTCACCATACTTGTCTCAATTATCATCATAGAAATCCA 1039
QY 582 -----TrrleArglleValPheLeuCysTrrleProTrrple 594
Db 1040 GATACGATGAATGAGTGAATGATAGATAATATTCCTTTATTTGTTACCTTGCGATA 1099
QY 595 LeuArgMetSerArgProGlyArgProLeulleleLeuGluPheProThrThrProCysSer 614
Db 1100 TTGGCATGCAAGACCCGACAG--GTTGGCTACGAATGTCGCGCCGCCCTCTTCT 1156
QY 615 AspThrSerSer-----GluArglyshHisGlnIleLeuSerAspValGluLeuLys 631
Db 1157 TCGAGTTCTCTCGCATCCGGCAGAGAGAACAGATC---CAAACGTTGAGCTCAAG 1213
QY 632 GluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHis 651
Db 1214 GAGAGGTCCTCCAAAGTCTCTGCTGCCAATGTGCTCGATATAGACGATGATTTCCGA-- 1270
QY 652 AsnCysArgProMetThrProGlyGlyThrLeuProHisasnProAlaPheTyrArgThr 671
Db 1271 ---TGCNAATCATGATGTGCCAGCGCATTTTGCCCCACCGCCACATATTACAGACG 1327
QY 672 ValTrrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThr----- 687
Db 1328 ATGTACAGGCAAGGGGATGACGCGAGCGTGGACCCGTCGGACGACGTCGTCAGTTGTG 1387
QY 688 -----ArgMetProAspAlaValThrHisThrCysIleLysSerSerThrGluTyr 705
Db 1388 GACGGGCGTTTGACGAGGCGCATTTCC---CACACCTGTCTGACATCCTCTGCGGAGTAC 1444
QY 706 GluLeuGlyLeulleleLeuLysGluIleAtgPheIleThrAspGlnLeuArgLysAspAsp 725
Db 1445 GAACTGGCGCTGATATCAAGGAGCTGCGTGTGATACAGACAGCTCAAAAAGAGGAC 1504
QY 726 GluCysAsnAspIleAlaAsnAspTrrLysPheAlaAlaMetValValAspArgLeuCys 745
Db 1505 GAAACAAGCGACATTAACGCGAGATTGGAATTTGCTGCCATGCTCGTCGATCGTTGTGC 1564
QY 746 LeullellePheThrMetPheAlalleLeuAlaThrIleAlaValLeuLeuSerAlaPro 765
Db 1565 CTTATTATTTTCACTTGTGTTTACTATTATATAGCAACCTGCTGCTACTTCTCAGCGCGG 1624
QY 766 HisIleIleValSer 770
Db 1625 CATTTCAATTGTGAGT 1639

RESULT 6
AF143846
LOCUS
DEFINITION
Heliothis virescens putative nicotinic acetylcholine receptor alpha
7-1 subunit mRNA, complete cds.
ACCESSION
AF143846
VERSION
AF143846.1 GI:4895004

```

KEYWORDS

SOURCE
ORGANISM

Heliothis virescens (tobacco budworm)
Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
1 (bases 1 to 3629)

REFERENCE

AUTHORS
TITLE

Schulte, T., Oellers, N. and Adamczewski, M.
Putative alpha subunits of insect nicotinic acetylcholine receptors
more similar to vertebrate alpha 7 subunits and C. elegans Ce21
than to other insect nicotinic acetylcholine receptor alpha
subunits
Unpublished

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

2 (bases 1 to 3629)
Schulte, T., Oellers, N. and Adamczewski, M.
Direct Submission
Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18,, Leverkusen
51368, Germany

FEATURES

source

Location/Qualifiers

1..3629

/organism="Heliothis virescens"

/mol type="mRNA"

/db xref="taxon:7102"

335..1825

/note="hvnachra7-1"

/codon start=1

/product="putative nicotinic acetylcholine receptor alpha
7-1 subunit"

/protein id="AA032697.1"

/db xref="GI:4895005"

/translation="MGRRARRSLAAPAGLLLLCLLMPRGACGCGYHEKRILLHLLDH
YNYLRFVNVNEDPLQLSGLITLQIIDVDKQLITNTLKLWMDMLRWNTSDP
GGVDRLPVRLKPDVLMYNGADGDFSTPTNVVVRNNGSLVYPRGIFKSTCKT
DIYFEDDORCBMKGSTYDGYQLDLQDEGGDISSFVTNGEWELIGVPGKRNE
IYVNCPEPYDITEAVVIRKLYLFFNLIVFCVLLASVALLGFTLTPDGGKLSLG
VTLLSLTVLNVMAETMPATSDAPLLGLYFNCIMFVMASSVVSTLILNLYHHRHAD
THEMSDRCVFLYLPWLVRMSRPGSATPPPARVPPDLELRSSKSLANVLD
IDDFRHPQAGCCQYRGGENGAGLAHSCFGVDYELSLILKEIRVITDQMRKD
DEDADISRDWKFAMVVDRLCLIFTLTIATLALLSAPHIMVS"

ORIGIN

Alignment Scores:

Pred. No.: 2,15e-125 Length: 3629
Score: 1850.00 Matches: 361
Percent Similarity: 75.78% Conservative: 30
Best Local Similarity: 69.96% Mismatches: 49
Query Match: 45.76% Indels: 76
DB: 3 Gaps: 7

US-09-303-232-2 (1-770) x AF143846 (1-3629)

QY 295 LeuLeullePyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArg 314
Db 383 CTGCTGCTGCTGCTGCTGCTGCTGCGCGAGGGGGCGCGCTGCGGGTACCACGAGAGCGG 442
QY 315 LeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSer 334
Db 443 CTACTGCACCACCTATTGGACCACCTACCAAGTACTGAGAGAGGCCCTCGTCAACGAGAC 502
QY 335 AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGluLys 354
Db 503 GACCGCTGCAGCTCTCTTCGCGCTCAGCTCATGCATCATCATCATCATCATCATCATCAT 562
QY 355 AsnGlnLeuLeuValThrAsnValTrrLeuLysLeuGluTrrAsnAspMetAsnLeuArg 374
Db 563 AACCGCTTTTAAATACAAACATCTGCTGCTAAACATAGAGTGAATGATGATGATGATGATG 622
QY 375 TrrAsnThrSerAspTrrGlyGlyValLysAspLeuArgIleProProHisArgIleTrr 394
Db 623 TGGAAACATCTCAGATTTCGGCGGGTCAAGATTTAAGAGTCCACCCACAGACTATGCG 682
QY 395 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr 414

QY 375 TrpAsnThrSerAspTyrGlyValLysAspLeuArgIleProProHisArgIleTrp 394
Db 623 TGGAAACATTTCAGATTTCGGGGGGTCAAGATTTAAGAGTGCCACCCACAGACTATGG 682
QY 395 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr 414
Db 683 AAACAGACGCTCTTATGTACAAACAGCGCGAGCAAGAGGTTTCGACAGACGATATCCACAG 742
QY 415 AsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer 434
Db 743 AACGTGGTGGTGGCGAACAACCGCTCGTGTCTGACGTGCGCGCGCATCTTCAAGAGC 802
QY 435 ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe 454
Db 803 ACCTGCAAGATCGACATCACTCGTTCCTCCCTTCACAGCACCAACGATGCGAGATGAAGTTT 862
QY 455 GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGluThrGlyGly 474
Db 863 GGCAGCTGGACTTATGATGTTTATCAGTTGGATCTACAACCTACAGGATGAAGGGGGCGGA 922
QY 475 AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArg 494
Db 923 GATATAAGAGTTTGTCAAGATGGCGAATGGAGTTAATAGGAGTCCCGCGCAAGCGC 982
QY 495 AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIle 514
Db 983 AACGAGATCTACTACAACTGTGTCCGAGGCCATACATCGACATCAGTTTTCGGTGGTG 1042
QY 515 IleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla 534
Db 1043 ATCCGAGGAAACGCTCTACTTCTTCAATCTGATCGGCCCTCGGTGCTCATCGCC 1102
QY 535 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly 554
Db 1103 TCAATGGCTCTATGGGGTTTCACTTCCCTCCAGACTCCGAGAGAAAGTTGCTTTAGGT 1162
QY 555 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 574
Db 1163 GTGACGATATTAATGTCGTGAGCGGTGTTCTCAATGTTGGCGGAGAGATGCCAGCG 1222
QY 575 ThrSerAspAlaValProLeu----- 581
Db 1223 ACCTCGGACGCGCTGCTTGTCTCGGCACCTTCACTCAATGTCATCATGTTTCATGGTGGCT 1282
QY 581 ----- 581
Db 1283 TCCTCGTCTCTCCACCTACTGATCTCAACTACCAACCGCGCAGCAGACACTCAC 1342
QY 582 -----TrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMet 597
Db 1343 GAAATGAGTGATGGATTCTGTCGTGTTCTTTATTTGGTTCGCGGGTGGCGGATG 1402
QY 598 SerArgProGly-----ArgProLeuIleLeuGluPheProThrThrProCysSer 614
Db 1403 TCACGCGCGGCTCGGCGAGCAGCGCGCGCGCGGTACCTCGCGCGCG----- 1456
QY 615 AspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSer 634
Db 1457 -----GACTGGAGTGGCGAGCGCTCC 1480
QY 635 SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArg 654
Db 1481 TCCAAGTCGCTCTAGGAAGCTGCTGACATCGATCGACTTCGCCAC----- 1531
QY 655 ProMetThrProGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly 674
Db 1532 -----CCGCAAGCGCAGCAGCGCAATGTCGCGTACTACAG----- 1570
QY 675 GlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThr 694
Db 1571 GGGGGTGAGAGAAATGCGCGGG-----TTGGCG 1600

QY 695 HisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeuLysGluIle 714
Db 1601 CGGCACAGATTCTTTC-----GGTGTCCAGCTACGAGCTCTCCCTCATTTCTGAAGAGATT 1654
QY 715 ArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrp 734
Db 1655 AGAGTCATCAGATCAGATCGCAAGACACAGAGATGCGGACATTTCCGCGACTGG 1714
QY 735 LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIle 754
Db 1715 AAGTTCCGCGCATGTTGTCGTCGACAGACTGTCCTTATTATCTTTACCTCTTCAATC 1774
QY 755 LeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIleValSer 770
Db 1775 ATCCCAAGCTAGCTGCTGCTCGCGCCACACATCATGTTGTGTCG 1822

RESULT 8
LOCUS E58347 3701 bp DNA linear PAT 18-JUN-2001
DEFINITION Nucleic acid encoding insect acyl choline receptor subunit.
ACCESSION E58347
VERSION E58347.1 GI:13019346
KEYWORDS JP 2000023680-A/2.
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3701)
AUTHORS Martin, A., Nadja, E. and Thomas, S.
TITLE Nucleic acid encoding insect acyl choline receptor subunit
JOURNAL Patent: JP 2000023680-A 2 25-JAN-2000;
BAYER AG
COMMENT OS Heliothis virescens
PN JP 2000023680-A/2
PD 25-JAN-2000
PF 26-APR-1999 JP 1999118159
PR 04-MAY-1998 DE 19819829.9
PI MARTIN ADAMUTSUEUSUKI,NADJA ERASU,THOMAS SCHULTE PC
C12N15/09,A01K67/033,C07K14/705,C07K16/28,C12N1/21,C12N5/10, PC
C12Q1/68.
PC G01N33/15,G01N33/50//(C12N1/21,C12R1:19),C12N15/00,C12N5/00 CC

FEATURES
source
1..3701
/organism="Heliothis virescens"
/mol_type="genomic DNA"
/db_xref="taxon:7102"

ORIGIN
Alignment Scores:
Pred. No.: 2,21e-125 Length: 3701
Score: 1850.00 Matches: 361
Percent Similarity: 75.78% Conservative: 30
Best Local Similarity: 69.96% Mismatches: 49
Query Match: 45.76% Indels: 76
DB: 6 Gaps: 7

US-09-303-232-2 (1-770) x E58347 (1-3701)

QY 295 LeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArg 314
Db 383 CTGCTGCTGCTGCTGCTGCTGCGCGAGGGGGGACGCTGCGGGTACCACGAGAGCGG 442
QY 315 LeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSer 334
Db 443 CTACTGCACCACTATTGGACCACTACAACGTAACGAGAGCCCGCTCGTCAACGAGAC 502
QY 335 AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleAspValAspGluLys 354
Db 503 GACCCGCTGACGCTCTCCCTTCGGCTCAGCTCATCGATCATCGTGCAGAGAG 562

```
Qy 355 AsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArg 374
Db 563 AACCCAGCTTTTAAATAACAAACATCTGGCTAAACATAGAGTGAATGATATGAACCTTGAAG 622
Qy 375 TrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProProHisArgIleTyr 394
Db 623 TGGAACACTTCAGATTTCGGCGGGGTCAAGATTTAAGATGCGCACCCACACAGACTATGG 682
Qy 395 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr 414
Db 683 AAACCCAGACGCTCCTTATGTACAAACAGCGCGGACGAGGGTTCGACAGCAGCTATCCAAACG 742
Qy 415 AsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer 434
Db 743 AACGTGGTGGTGGGAAACACCGCTCGTGTCTGTACGTGGCGCCCGCGCACTCTCAAGAGC 802
Qy 435 ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe 454
Db 803 ACCTGCAAGATCGACATCACCTGGTTCCCTTTCGACGACCAACAGATCGAGATGAAGTTT 862
Qy 455 GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGly 474
Db 863 GGCAGCTGGACTTATGATGGTTATCAGTTGGATCTACAACCTACAGGATGAAGGGGGCGGA 922
Qy 475 AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArg 494
Db 923 GATATAAGCAGTTTGTTCACGAATGGCGAATGGGAGTTAATAGGATGCCCGCGGACGCCG 982
Qy 495 AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIle 514
Db 983 AACGAGATCTACTACACTGTTGTCGGAGCCATACATCGACATCAGCTTTGCGGTGGT 1042
Qy 515 IleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla 534
Db 1043 ATCCGGAGAAACGCTACTACTTCTTCAATCTGATCGTGCCTCGCTCATCGCC 1102
Qy 535 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly 554
Db 1103 TCCATAGGCTCTATTGGGGTTTACCTTGCCTCCAGACTCCGGAGAAAGTTGCTTTAGGT 1162
Qy 555 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 574
Db 1163 GTGAGGATATTACTGCTGTGACGGTGTCTCAACATGGTGGCGGACGATGCCAGCG 1222
Qy 575 ThrSerAspAlaValProLeu----- 581
Db 1223 ACGTCGGACGCGCTGCCCTTGTCTCGGCACCTACTTCAACTGCATCATGTTTCATGGTGCT 1282
Qy 581 ----- 581
Db 1283 TCCTCCGTGCTCCACCATACTGATCTCTCACTACCAACCACCGGCGACGACACTCAC 1342
Qy 582 -----TrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMet 597
Db 1343 GAAATGAGTGTGATTCGTTGCTGTTGCTTTTATTTGGTCTCGCTGGTGTGCGCATG 1402
Qy 598 SerArgProGly-----ArgProIleLeuLeuGluPheProThrThrProCysSer 614
Db 1403 TCACGCGCGGCTCGCGGACGACGCGCGCGCGCGCGCTACCTCCGCGCGCG----- 1456
Qy 615 AspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSer 634
Db 1457 -----GACCTGGAGCTCGCGAGCGCTCC 1480
Qy 635 SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHisAsnCysArg 654
Db 1481 TCCAAAGTCGCTCCTAGCGAAACGCTGCTGACATCGATGACGACTTCGCGCAC----- 1531
Qy 655 ProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly 674
Db 1532 -----CCGCAAGCGGACGAGCGGCAATGCTGCGGATACTACAGG----- 1570
```

```
Qy 675 GlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThr 694
Db 1571 GCGGGTGGAGAGATGGCGGG-----TTGGCG 1600
Qy 695 HisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeuLysGluIle 714
Db 1601 GCGCACAGTTGCTTC-----GGTGTGCTAGCTACGAGCTCTCCCTCATCTTGAAGGAGATT 1654
Qy 715 ArgPheIleThrAspGlnLeuArgLysAspGluCysAsnAspIleAlaAsnAspTrp 734
Db 1655 AGAGTCATACAGATCAGATCGCGCAGGACGACGAGATGGGACATTCGCGGAGCTGG 1714
Qy 735 LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIle 754
Db 1715 AAGTTGCGCGCCATGGTGGTGGACAGACTGTGCTTATTATCTTACCTGTTCCACAATC 1774
Qy 755 LeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIleValSer 770
Db 1775 ATCCGCCGCTAGCGTGTGCTGTCGCGCCACACATCATGTGGTGG 1822
RESULT 9
AF321446 2023 bp mRNA linear INV 29-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6
DEFINITION subunit variant type II (nAChRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION AF321446
VERSION AF321446
KEYWORDS AF321446.1 GI:20152846
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2023)
AUTHORS Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
TITLE Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
JOURNAL Genetics 160 (4), 1519-1533 (2002)
MEDLINE 21969411
PUBMED 11973307
REFERENCE 2 (bases 1 to 2023)
AUTHORS Grauso, M. and Sattelle, D.B.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK
FEATURES
source
1. 2023
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1. 2023
/gene="nAChRalpha-30D"
379..1863
/gene="nAChRalpha-30D"
/feature="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3b and 8a"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit
variant type II"
/protein_id="AA013393.1"
/db_xref="GI:20152847"
/translation="MDSPLASLSFVLLIFLAIKESQCGPHEKRLNHLHLLSTNTYL
ERPVAESEPLEVFGTLQOIIDVDEKNQLLITNLISLEWNLNRLNWEYGGVK
DLRITPNKLWKPDLVYNSADGFTVHTNLVKGSGSLYVPGIIFKSTCKMDITW
PPFDHOCMEKFGSWTYDGNQLDVLNSEDGGLSDFTNGEWYLLAMPKKNITVYA
CCPEPYDITFTIQRRLTYFFNLIVPCVLISSMALLGFTLPDPSGKLTGLVTIL
```

LSLTVFLNLVAESMPTSDAVPLIGTYFNCIMFVASSVLTIVVVLNVHRTADHEM
 PPWTKSFLQWLPWILNRMPGRKTRKTIILSNRMKELEKSESSKSLANVLDDIDD
 DFRTISQSTALSSSPFRPTVEHRTAIGCNKHDLKELHLELOFITARMKRADD
 EAEILGWKFAAMVDRLFLVTLFTIIATVTVLLSAPHIIVQ

variation

/gene="nAcRalpha-30D"
 /note="results in glycine to asparagine substitution;
 compared to the sequence deposited in GenBank Accession
 Number AE003626"
 /replace="aa"
 794

variation

/gene="nAcRalpha-30D"
 /note="results in glycine to serine substitution; compared
 to variant clone"
 /replace="a"
 846

variation

/gene="nAcRalpha-30D"
 /note="results in asparagine to serine substitution;
 compared to the sequence deposited in GenBank Accession
 Number AE003626 and in variant clone"
 /replace="a"
 1314

variation

/gene="nAcRalpha-30D"
 /note="compared to variant clone"
 /replace="a"
 1560

variation

/gene="nAcRalpha-30D"
 /note="compared to variant clone"
 /replace="t"
 1683

variation

/gene="nAcRalpha-30D"
 /note="compared to variant clone"
 /replace="t"
 1683

ORIGIN

Alignment Scores:

Pred. No.:	2,56e-109	Length:	2023
Score:	1625.50	Matches:	320
Percent Similarity:	70.44%	Conservative:	47
Best Local Similarity:	61.42%	Mismatches:	75
Query Match:	40.21%	Indels:	79
DB:	3	Gaps:	7

US-09-303-232-2 (1-770) x AF321446 (1-2023)

QY	291	SerTrpIlePheLeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr	310
Db	406	TCGCTGTTTCTCTGTGTGATCTTTCTGGCGATAATTAAGAAAGCTGT---	CAAGGACCT 462
QY	311	HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal	330
Db	463	CATGAAAGCGCTCTGTAACCATCTCTGTCCACCTACAAATAGCTGGAGCGACCGGTG	522
QY	331	LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAsp	350
Db	523	GCCAAATGAATCGGAGCCCTCGAGTTAAGTTCGACTGACGCTGCAGAGATCATCGAC	582
QY	351	ValAspGluLysAsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp	370
Db	583	GTGCGAGAGAGAAATCAGCTTCTCATACCAATCTTTGGCTTTGGTGGAGTGAACGAC	642
QY	371	MetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProPro	390
Db	643	TACAACTCGCTGGATGAACAGGAATACGGCGGGTCAAGGATCTACGAATCAGCCCC	702
QY	391	HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly	410
Db	703	AACAAGCTGTGGAAGCCGACGCTCATGTACAAACAGCGCGGATGAGGATTCGATGGC	762
QY	411	ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGly	430
Db	763	ACGTATACACCAACATTTGTGTGTCAAACATCGGCGGCGAGTTGTCTGTACGTGCCCTTGT	822

QY	431	IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys	450
Db	823	ATCTTCAAGAGCACATCAAGATGACATCACTCGTGGTTCCTCATTTGATGACCAACATTGC	882
QY	451	GluMetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAsp	470
Db	883	GAATGAATTCGGTAGTTGGACTTACATGGAATCAGTTGGATTTTGGTTTGAATTC	942
QY	471	GluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuGlyVal	490
Db	943	GAAGATGGAGGGGATCTTCCGATTTTCATAACAATGGCGAGTGGTACTTCTGCTGCCATG	1002
QY	491	ProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr	510
Db	1003	CCGGAAAGAAAGAAATACGATAGTCTAGCCCTGCTGCCAGAACCATATGTGATATCACC	1062
QY	511	PheAlaIleIleIleArgArgThrLeuTyrTrpPhePheAsnLeuIleIleProCys	530
Db	1063	TTTACTATACAAATTCGTCGCCGTACATTAATATTTTTCAAATTAATCGTCCCATGT	1122
QY	531	ValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLys	550
Db	1123	GTGCTAATCTCATCGATGGCCCTACTGGCTTCACATTCGCCCGGATTCGGCGAGAAA	1182
QY	551	LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGlu	570
Db	1183	CTGACGCTGGCGGTAACTACTACTACTATTAACAGTATTTCTTAAACCTTGTGCGCGAG	1242
QY	571	ThrMetProAlaThrSerAspAlaValProLeu	581
Db	1243	TCCATGCCGACACGTCGGATGCTGTCTCTTATAGGACCTACTTCAATTGGCATGT	1302
QY	581	-----	581
Db	1303	TTCATGTGCTGCTCGTGGTGGTGTGACAGTAGTGGTCTCAACTACCACCATCGCACA	1362
QY	582	-----	582
Db	1363	GCGGACATTCACGAGATGCCACCGTGGATCAAGTCCGTTTCTTACAAATGGTCCCTGG	1422
QY	594	IleLeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrProCys	613
Db	1423	ATCTTGGCAATGGTGGACCCGGTCGCAAGAT	1455
QY	614	SerAspThrSerSerGluArgLysHisGlnIleLeuSer	629
Db	1456	-----ACACGCAAAACAATACTATTAAGCAATCGCATGAAGAGCTGGAG	1500
QY	630	LeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPhe	649
Db	1501	CTAAAGAGCGCTCTCCAAATCCCTGTCGCCAATGTCCTCGACATCGACGACGACTTC	1560
QY	650	ArgHisAsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyr	669
Db	1561	CGGCACACAATA	1572
QY	670	ArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMet	689
Db	1573	-----TCTGGCTCCCAACCGCCATTTGGCTCGTGGCCAGCTTCGGTGG	1617
QY	690	ProAspAlaVal	708
Db	1618	CCACAAACGGTGGAGGAGCATCACAGGCCATC---GGTGTCAATCACAAGATCTTCAT	1674
QY	709	LeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsn	728
Db	1675	CTAATTTCTCAAGAATTCGAATTTATACGCCCGGATGCGCAAGACTGACGACGAGCG	1734
QY	729	AspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIle	748
Db	1735	GAATTCATCGGCGATTGGAAGTTTCGCGCAATGTTGTGATAGATTTTGTAAATGTT	1794
QY	749	PheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIle	768


```
Db 943 GAACATGAGGGGATCTTCCGATTTCATAACAATGCGAGTGTGATCTCTTCCATG 1002
QY 491 ProGlyysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr 510
Db 1003 CCGGAAAGAAAGATAAGATAGTACGCTGCTGCCAGAACCATATGTGATATCACC 1062
QY 511 PheAlaIleIleLeuArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCys 530
Db 1063 TTACTATACAAATCGTCGCGTACATATATATTTTCAATTAATCGTCCATGT 1122
QY 531 ValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLys 550
Db 1123 GTGCTAATCTCATCGATGGCCCTACTGGGCTTCACATTGCCGCGGATTCGGCGAGAA 1182
QY 551 LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGlu 570
Db 1183 CTGACGCTGGAGTTCAATTCCTATCGCTCACAGTGTTCCTCAACCTTGTAGCTGAG 1242
QY 571 ThrMetProAlaThrSerAspAlaValProLeu----- 581
Db 1243 ACATTGCCCCCAAGTATCTGATGCAATCCCTTGTAGCACCTACTTCAATTGATCATG 1302
QY 581 ----- 581
Db 1303 TTCATGTCGCTCGTGGTGGTGTGACAGTAGTGTCAACTACCACCATCGCACA 1362
QY 582 -----TTPileArgIleValPheLeuCysTyrTrpLeuProTyr 593
Db 1363 GCGACATTCACGAGATGCCCGGTGATCAAGTCCGTTTCTTACAAATGGCTGCCCTGG 1422
QY 594 IleLeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrThrProCys 613
Db 1423 ATCTTGCATGGTGCACCGCTCGCAAGT----- 1455
QY 614 SerAspThrSerSerGluArgLysHisGlnIleLeuSer-----AspValGlu 629
Db 1456 -----ACACGCAAAACAATACTATTAAAGCAATCGCATGAAGAGCTGGAG 1500
QY 630 LeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAsp 649
Db 1501 CTAAGAGCGCTCTCCAAATCCCTGCTGGCCAATGCTCTGCACATCGACGACGACTTC 1560
QY 650 ArgHisAsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyr 669
Db 1561 CGGCACACAATA----- 1572
QY 670 ArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMet 689
Db 1573 -----TCTGGCTCCCAACCGCATTTGGCTCGTCCGCCAGCTTCGGTCGG 1617
QY 690 ProAspAlaVal---ThrHisHisThrCysIleIleLysSerSerThrGluTyrGluLeuGly 708
Db 1618 CCACACAGGTGAGGAGCATCACGCGCATC---GGCTGCAATCACAAAGATCTTCAT 1674
QY 709 LeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsn 728
Db 1675 CTAATCTCAAGAATTGCAATTATTACGCGCGGATCGCAAGCTGACGACGAGCG 1734
QY 729 AspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIle 748
Db 1735 GAATTGATCGCGCATTTGGAAGTTCGCGGCAATGGTTGTGATAGATTTTGTATTAATGTT 1794
QY 749 PheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIle 768
Db 1795 TTCACGCTCTTCAGTATTATTGCAACGGTTACGGTGTCTCTCGCTCCGCTCGCACAATA 1854
QY 769 Val 769
Db 1855 GTG 1857
DME554209 1665 bp mRNA linear INV 01-APR-2003
LOCUS
```

```
DEFINITION Drosophila melanogaster mRNA for nicotinic acetylcholine receptor
subunit Dalpha6 (nAcRalpha-30D gene).
ACCESSION AJ554209
VERSION 1
KEYWORDS nAcRalpha-30D gene; nicotinic acetylcholine receptor subunit
Dalpha6.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1665)
AUTHORS Millar,N.S.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) Millar N.S., Department of Pharmacology,
University College London, Gower Street, London, WC1E 6BT, UNITED
KINGDOM
FEATURES
source
1..1665
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1..1665
/gene="nAcRalpha-30D"
44..1528
/gene="nAcRalpha-30D"
/codon_start=1
/product="nicotinic acetylcholine receptor subunit
Dalpha6"
/protein_id="CAD86935.1"
/db_xref="GI:29466435"
/translation="MDSPLPASLSFLVLIIFLAIKESQGPHEKRLNHLSTNTL
ERVANESPELVKFTLQIITDDEKNQLITNLWLSLEWMDYLRNMTETGGVK
DLRITENKLMKPDVLMYNSADEGDTYHTSVVVGSGCLYVPGFKSTCKMDIYW
FPDDQCHCKMKFGSWTDGNQLDLVLSDEGDLSDPITNGEWVLLAMPKKNITVA
CPEPYVDITFTIQRRTLYFPNLIIVPCVLSSMALLGFTLPPSGEKLTLGVTIL
LSITVNLVAESMTTSDAPLIGTIFNCIMFMAVSSVLTIVVLYNLYHRTADIHEM
PMTKSVFLQWLPWILLRMGRPGKTRKTIILSNRMKELEKERSKSLIANVLDIDD
DPHHTIGSQTAIGSASFRTTVEEHHTAIGCNHKLHLILKELQFIETMRKADD
EALIGDWKFAAMVVDVRFCLIVFTLTITATVTLLSAPHIIQV"
44..121
sig_peptide
/gene="nAcRalpha-30D"
122..1525
mat_peptide
/gene="nAcRalpha-30D"
/product="nicotinic acetylcholine receptor subunit
Dalpha6"
ORIGIN
Alignment Scores:
Pred. No.: 3,89e-109 Length: 1665
Score: 1621.50 Matches: 320
Percent Similarity: 70.44% Conservative: 47
Best Local Similarity: 61.42% Mismatches: 75
Query Match: 40.11% Indels: 79
DB: 3 Gaps: 7
US-09-303-232-2 (1-770) x DME554209 (1-1665)
QY 291 SerTrpIlePheLeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr 310
Db 71 TCGCTGTTTGTCTGTGTGATCTTTCGGGATTAATTAAGAAAGCTGT---CAAGACCT 127
QY 311 HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal 330
Db 128 CATGAAAGCGCTGCTGAACCATCTGCTGCCACCTACATACGCTGGAGCGACCCGTG 187
QY 331 LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAsp 350
Db 188 GCCAATGAATCGGAGCCCTCGGAGTTAAGTTCCGACTCAGCTGCGAGCATCATCGAC 247
QY 351 ValAspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp 370
Db 248 GTGGACGAGAGAAATCAGCTTCTCATACGAATCTTTGGCTTTCGTTGGATGGACGAC 307
```

```

Qy 371 MetAsnLeuArgTrrAsnThrSerAspTyrGlyValVallyAspLeuAqIleProPro 390
Db 308 TACAACTCGCTGGATGAACCGAATACGGCGGGTCAAGGATCTACGATCACGCC 367
Qy 391 HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410
Db 368 AACAAAGCTGTGGAAGCCGACGCTGCTCATGTACAAACAGCGCGGATGAGGATTCGATGCC 427
Qy 411 ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGly 430
Db 428 ACGTATCACACCGAGCTGTGGTCAAAATCGCGCGAGTGTCTGTACGTGCCCCCTGGT 487
Qy 431 IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys 450
Db 488 ATCTTCAGAGCACATGCAAGATGACATACAGTGGTTCCTCATTTGATGACCAATGTC 547
Qy 451 GluMetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAsp 470
Db 548 GAAATGAAATTCGGTAGTGGACTTACGATGAAATCAGTTGGATTTGGTTTGAGTTCC 607
Qy 471 GluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrrPheLeuLeuGlyVal 490
Db 608 GAAGATGAGGAGGATCTTTCCGATTTTCATAAATGGCAGTGGTACTTGTCTGGCCATG 667
Qy 491 ProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr 510
Db 668 CCGGGAAGAGATACATAGTCTACGCTGCTGCCAGAACCATATGTCATATCACC 727
Qy 511 PheAlaIleIleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCys 530
Db 728 TTTACTATACAAATTCGTCGCGGTACATATATATATTTTCAATTTAATTTGGCCATGT 787
Qy 531 ValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlyLys 550
Db 788 GTGCTAATCTCATGATGGCCCTACTGGGCTTTCATTCATTCGCGCGGATTCGGCGAGAAA 847
Qy 551 LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGlu 570
Db 848 CTGACGCTGGCGGTAACTATCTACTATCAATACAGATATTTCAAACTTGTCCGCGAG 907
Qy 571 ThrMetProAlaThrSerAspAlaValProLeu----- 581
Db 908 TCCATGCCGACACGTCGATGCTGTCTCTCTTATAGGACACTTCAATTTGCATCATG 967
Qy 581 ----- 581
Db 968 TTCATGTGCGCTCGTCGCTGGTGTGACAGTAGTGTGCTCACTACACATCGCACA 1027
Qy 582 -----TrpIleArgIleValPheLeuCysTrpLeuProTrp 593
Db 1028 GCGGACATTCACGATGCCACCGTGGATCAAGTCGTTTCTTCAATAGTGGTGGCCCTGG 1087
Qy 594 IleLeuArgMetSerArgProGlyArgProIleLeuLeuGluPheProThrThrProCys 613
Db 1088 ATCTTGCAGATGGTTCGACCGCGTGCAGAT----- 1120
Qy 614 SerAspThrSerSerGluArgGlyHisGlnIleLeuSer-----AspValGlu 629
Db 1121 -----ACGCGAAACAAATCTATTAAAGCAATCGCATGAAGAGCTGGAG 1165
Qy 630 LeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPhe 649
Db 1166 CTAAAGGAGCGCTCTCCAAATCCCTGTGGCCAAATGCTCTGACATCGACGACGACTTC 1225
Qy 650 ArgHisAsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyr 669
Db 1226 CGGCACACAATA----- 1237
Qy 670 ArgThrValTyrGlyGlnGlyAspGlySerIleGlyProIleGlySerThrArgMet 689
Db 1238 -----TCTGGCTCCCAACCGCCATTTGCTCGTCGCCGACGCTTCGTCGG 1282

```

690 ProAspAlaVal---ThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGly 708
1283 CCCAACCGCTGGAGGAGCATCACACGCCCATC---GGCTGCAATCAAAAGATCTTCAT 1339
709 LeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsn 728
1340 CTAATTCCTAAAGAAATTCGAATTTATTACGCGCGGATGCGCAAGCTGACGACGAAGCG 1399
729 AspIleAlaAsnAspTrpLysPheAlaMetValValAspArgLeuCysLeuIleIle 748
1400 GAATGATCGCGATTCGAAGATTCGCGCAATGTTGGATAGATTTGTTAATGTT 1459
749 PheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIle 768
1460 TTCACGCTCTTCACGATTTATTCACCGTTACCGTGTGCTCTCCGTCGCCACATATC 1519
769 Val 769
1520 GTG 1522

RESULT 12
LOCUS BT011147 1699 bp mRNA linear INV 19-DEC-2003
DEFINITION Drosophila melanogaster GH1518 full insert cDNA.
ACCESSION BT011147
VERSION BT011147.1 GI:40216007
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1699)
Stapleton, M., Brokstein, P., Hong, L., Agbavani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Friese, E.,
George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G.,
Miranda, A., Mungall, C. J., Nunoo, J., Paclell, J., Paragas, V., Park, S.,
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M.
and Celisnik, S.
Direct Submission
Submitted (19-DEC-2003) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcriptase errors that result in single base changes.
For further information about this sequence, please visit our location
and relationship to other sequences, please visit our Web site.
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.
location/Qualifiers
1. 1699
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
<1. 1699
/gene="ncAlpha-30D"
/db_xref="FLYBASE:FBgn0032151"
<1. 1528
/gene="ncAlpha-30D"

FEATURES
source
gene
CDS


```

QY      630  LeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPhe 649
Db      1501 CTAAGAGCGCTCTCTCAAAATCCCTGCTGCCAATGTCTTCGACATCGACGACGACTTC 1560
QY      650  ArgHisAsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyr 669
Db      1561 CGGCACACAATA-----1572
QY      670  ArgThrValTyrGlyGlnGlyAspAspCysSerIleGlyProIleGlySerThrArgMet 689
Db      1573 -----TCTGGCTCCCAACCCCAATGCTGCTCGGCAGCTTCGTCGG 1617
QY      690  ProAspAlaVal-----ThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGly 708
Db      1618 CCCACACGGTGGAGGAGCATCACACGCCATC---GGCTCAATCAAAAGATCTTCAT 1674
QY      709  LeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsn 728
Db      1675 CTAATTCCTCAAGAAATGCAATTTATTACGGCGCGGATGCGCAAGCTGACGACGAGCG 1734
QY      729  AspIleAlaAsnAspTyrLysPheAlaMetValValAspArgLeuCysLeuIleIle 748
Db      1735 GAATGTGATCGCGGATGGAAGTTCGCGGCAATGCTGCGGATAGATTTGTTAAATGTT 1794
QY      749  PheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIle 768
Db      1795 TTCACGCTCTTCAGATATTGCAACGGTACGCTGCTCTCCGCTCCGCATATC 1854
QY      769  Val 769
Db      1855 GTG 1857

RESULT 14
LOCUS   AF321449
DEFINITION Drosophila melanogaster 2110 bp mRNA linear INV 29-APR-2002
          subunit variant type V (nAcRalpha-30D) mRNA, complete cds,
          alternatively spliced.
ACCESSION AF321449
VERSION   AF321449.1 GI:20152852
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE   Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Phylodidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2110)
          Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
          Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
          Dalphea5, Dalphea6 and Dalphea7, in Drosophila melanogaster Identify a
          New and Highly Conserved Target of Adenosine Deaminase Acting on
          RNA-Mediated A-to-I Pre-mRNA Editing
          Genetics 160 (4), 1519-1533 (2002)
JOURNAL  MEDLINE
          PUBMED 21969411
          11973307
REFERENCE 2 (bases 1 to 2110)
          Grauso, M. and Sattelle, D.B.
          Direct Submission
          Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
          Genetics Unit, University of Oxford, South Park Road, Oxford OX1
          3QX, UK
FEATURES             Location/Qualifiers
     source           1..2110
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /chromosome="2"
                     /map="30D1"
                     /dev_stage="embryo"
     gene             1..2110
                     /gene="nAcRalpha-30D"
     CDS               379..1950

```

```

/gene="nAcRalpha-30D"
/notes="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3a, 8a and
8b; contains two repeats of the transmembrane TM2 region"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalphea6 subunit
variant type V"
/protein_id="AAM13396.1"
/db_xref="GI:20152853"
/translation="MDSPLPASISLFVLLIFLAIKESQCPHEKRLINHLSTYNTIL
ERPVAESEPLEVKFGLTLOQIIDVDEKNOILITNAMLEWNYDLNRWNETEYGVK
DLRTTNKLMKPDVLMYNSADEGDTYHTNIVVKNHNGSLYVPGGFKSKCKMDITW
FPDDDOCEMKFGSWYDGNQLDLVNSEGGDLSDFITNGEWLLAMPKKNITIVIA
CCPEYVDITFTTQIRRTTYFFNLIVPCVLISSMALLGFTLPDPSGEKLTGLVTIL
LSUTVFNLUVAESMPTSDAVPLIGVLTLLSLTVFLNVAETLPQVSDAIPLLGTYN
CIMFVASSVLTVVNLVNHRTADIHMPFWIKSVFLOLWPLILMRGRPKIRTKT
IILSNRMKEHLKERSKSLANVLDDDDFRHTISGQTAIGSSASFGPTTVEREH
TAIGCKNHKDLHLIKELQFITARMRRADDEAELIGDKFAMVVDFCLVFLFTIL
ATVTLLSAHHIIVQ"
846
/variation
/gene="nAcRalpha-30D"
/notes="results in methionine to isoleucine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"
1401
/variation
/gene="nAcRalpha-30D"
/notes="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
1647
/variation
/gene="nAcRalpha-30D"
/notes="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
1770
/variation
/gene="nAcRalpha-30D"
/notes="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"

ORIGIN
Alignment Scores:
Pred. No.:      3,11e-108      Length:      2110
Score:          1611.00      Matches:      320
Percent Similarity: 66.55%      Conservative: 46
Best Local Similarity: 58.18%      Mismatches:   76
Query Match:    39.85%      Indels:       108
DB:              3              Gaps:         7

US-09-303-232-2 (1-770) x AF321449 (1-2110)
QY      291  SerTrpIlePheLeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr 310
Db      406  TCGCTGTTCTCTCTGTTGATCTTTCTGGCATATTAAGAAAGCTGT---CAAGACCT 462
QY      311  HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal 330
Db      463  CATGAAAAGCGCTGCTGGAACCATCTGCTGCCACCTACAAATACGCTGGAGCGACCGTG 522
QY      331  LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIleAsp 350
Db      523  GCCAATGAATCGAGCCCTCGAGGTAAAGTTGGACTGACGCTGCGAGCATCATCGAC 582
QY      351  ValAspGluLysAsnGlnLeuValThrAsnValTyrLeuLysLeuGluTyrAsnAsp 370
Db      583  GTGGATGAAAGATCATAGATTCTGCACCAAAATCGCTGGTTAAATTTGAGTGAACGAC 642
QY      371  MetAsnLeuArgTrpAsnThrSerAspTyrGlyValLysAspLeuArgIleProPro 390
Db      643  TACAATCTCGCTGGAATGAAACGGAATACGCGGGGTCAAGGATCTACGAATCACGCC 702
QY      391  HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410

```

```

Db 703 AACAGCTGTGGAGCCCGAGTCTCATGTACACGCGGATGAGGATTCGATGC 762
Qy 411 ThrTyrGlnThrAsnValValArgAsnGlySerCysLeuTyrValProGly 430
Db 763 ACGTATCACACCAATGTGGTCAAAATAACCGCAGTGTCTGTACGTGCGCCCTGGT 822
Qy 431 IlePheLysSerThrCysLysIleAspIleThrTyrPheProPheAspAspGlnArgCys 450
Db 823 ATCTTCAAGAGCACATGCAAGATGGACATCACGTGGTTCCTTGTGATGACCAATTC 882
Qy 451 GluMetLysPheGlySerThrTyrAspGlyPheGlnLeuAspLeuGlnAsp 470
Db 883 GAATGAATTCGAGTGGACTACGATGAATCAGTTGATTTGGTTTGAATTC 942
Qy 471 GluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTyrGluLeuGlyVal 490
Db 943 GAAGATGGAGGGATCTTCCGATTTCATAACAATGGCGAGTGGTACTTGTTCGCCATG 1002
Qy 491 ProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr 510
Db 1003 CCGGAAAGAAATAGATAGTCTACGCTGCTGCCAGAACCATATGTCGATATACC 1062
Qy 511 PheAlaIleIleIleArgArgThrLeuTyrTyrPheAsnLeuIleIleProCys 530
Db 1063 TTTACTATACAAATTCGCGCGTACATATATATATTTTCAATTTAATCGTCCCATGT 1122
Qy 531 ValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLys 550
Db 1123 GTGCTAATCATCGATGGCCCTACTGGGCTTCATTTGCCCGCGGATTCGGGCGAGAA 1182
Qy 551 LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGlu 570
Db 1183 CTGACGTGGCGTAACTACTACTATCATTAACAGTATTTCTAATCTTGTGCGGAG 1242
Qy 571 ThrMetProAlaThrSerAspAlaValProLeu 581
Db 1243 TCCATGCCGACACGTCGGATGCTTCTCTTATAGGATTACAATTCCTTCTATCGCTC 1302
Qy 581 581
Db 1303 ACAGTGTCTTCAACCTTTAGCTAGCTAGACATTCGCCCAAGTATCTGATGCAATCCCTTG 1362
Qy 581 581
Db 1363 TTAGCACCTACTCAATTCATGTCATGTTCTATGTCGTCATCGTCGGTCTGCACAGTA 1422
Qy 582 582
Db 1423 GTGGTGTCAACTACCACTCGCACAGCGGACATTCACGAGATGCCACCGTGGATCAAG 1482
Qy 585 IleValPheLeuCysTyrLeuProTyrIleLeuArgMetSerArgProGlyArgProLeu 604
Db 1483 TCCGTTTTCTTACAAATGGCTCCCTGGATCTTGGCAATGGGTGCGACCGCGTCGAAGAT 1542
Qy 605 IleLeuGluPheProThrThrProCysSerAspThrSerSerGluArgLysHisGlnIle 624
Db 1543 1543
Qy 625 LeuSer 625
Db 1561 TTAAGCAATCGCATGAAGGAGCTGGAGCTAAAGAGCGCTCTCTCCAAATCCCTGTCGGCC 1620
Qy 641 AsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThrProGlyGly 660
Db 1621 AATGTCCTCGACATCGAGAGACTTTCGGCACAAATA 1659
Qy 661 ThrLeuProHisAsnProAlaPheTyrArgThrValTyrGlyGlnGlyAspGlySer 680
Db 1660 1660
Qy 681 IleGlyProIleGlySerThrArgMetProAspAlaVal---ThrHisHisThrCysIle 699

```

```

Db 1678 ATTGGCTCGTCGCGCAGCTTCGGTCGCCCCACACACGCTGGAGGAGCATCACACGCCCATC 1737
Qy 700 LysSerSerThrGluTyrGluLeuGlyLeuIleLysGluIleArgPheIleThrAsp 719
Db 1738 ---GGCTGCAATCACAAGATCTTCATCTAATTTCTTAAGAATTCGAATTTATTACGGCG 1794
Qy 720 GlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMet 739
Db 1795 CGGATCGCAAGCTGACGACGAGCGAATTCATCGCGATTCGGAAGTTTCGGCGCAATG 1854
Qy 740 ValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAla 759
Db 1855 GTTGTGATAGATTTTGTAAATGTTTTCACGCTCTTCAGATATTATTCGAACGGTTACG 1914
Qy 760 ValLeuLeuSerAlaProHisIleIleVal 769
Db 1915 GTGCTGCTCTCCGCTCCGCACATATCGTG 1944

RESULT 15
AF143847 3029 bp mRNA linear INV 27-MAY-1999
LOCUS Heliothis virescens putative nicotinic acetylcholine receptor alpha
DEFINITION 7-2 subunit mRNA, complete cds.
ACCESSION AF143847
VERSION AF143847.1 GI:4895006
KEYWORDS Heliothis virescens (tobacco budworm)
SOURCE Heliothis virescens
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE
1 (bases 1 to 3029)
Schulte, T., Oellers, N. and Adamczewski, M.
Purative alpha subunits of insect nicotinic acetylcholine receptors
more similar to vertebrate alpha 7 subunits and C. elegans Ce21
than to other insect nicotinic acetylcholine receptor alpha
subunits
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3029)
Schulte, T., Oellers, N. and Adamczewski, M.
Direct Submission
JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18,, Leverkusen
51368, Germany
FEATURES
Location/Qualifiers
source 1..3029
/organism="Heliothis virescens"
/mol_type="mRNA"
/db_xref="taxon:7102"
95..1600
/note="hvnachra7-2"
/codon_start=1
/product="putative nicotinic acetylcholine receptor alpha
7-2 subunit"
/protein_id="AAD32698.1"
/db_xref="GI:4895007"
CDS
/translaton="MAPMALALALLPVSEQEPHEKRLNALLANNYTLERPVANE
SEPLEVFLVLTQOIIDVENQQLITINILSWNDYNLWNSDEYGVKDLRITEN
KLWKPDLVFNYSDEGFDGTQTVNVRSGGSLYVPPGIFKSTCKMDIAWPFDDQH
CDMKPGSWTDGNDQLDLVKDEAGDLSDFITNGEWYLIQMGKKNTITYACCPYV
DVITFTIRRTLIYFNLIYFVCLISSMALLGWTLPDSEKLTGLVTLISLTVFL
NLVATETLPQVSDAIPLLGTIVFNCIMFVWASVLTIVVNLVHRTADIHMPQWIKSV
FLOWLPMILKRSRFGKTRKTIMNTRMRELEKRSKSLANVLIDIDDFRGGPP
PPNSTAGTGNLPGCSIPTDFRFRSPRPTMEDVGGGLGSHHREHLILRELOPITA
RMKKADEAEELISDWKFAAMVVDVDFCLVFTLTFTIATVAVLLSAPHIIVQ"
ORIGIN
Alignment Scores:
Pred. No.: 7.01e-108 Length: 3029
Score: 1609.00 Matches: 319
Percent Similarity: 71.62% Conservative: 57
Best Local Similarity: 60.76% Mismatches: 65
Query Match: 39.80% Indels: 84
DB: 3 Gaps: 10

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:39:04 ; Search time 808.783 Seconds
(without alignments)
4044.486 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKLTVDDELMLAV.....MFAILATIAVLISAPHIIVS 770

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09303232/runat_07052004_101109_23873/app_query.fasta_1.2261
-DB=N_Geneseq_29Jan04 -QFW=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232 @CGN 1.1200 @runat_07052004_101109_23873 -NCFU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4043	100.0	2886	3	Aaz24475
2	1850	45.8	3700	3	Aaz24476
3	1609	39.8	3109	3	Aaz24477
4	1498.5	37.1	984	4	ABL27131
5	1466.5	36.3	1540	4	ABL07231
6	1456.5	36.0	3144	4	ABL27130
7	1407.5	34.8	936	4	ABL13733
8	1253	31.0	803	4	ABL07799

9	1179.5	29.2	2769	2	AAT59196	Neuronal
10	1170.5	29.0	2101	2	AAT59197	Neuronal
11	1159.5	28.7	1509	3	AAC58395	Human PRO
12	1159.5	28.7	1509	4	AAC90380	Wild-type
13	1159.5	28.7	1876	2	AAV12197	Human neu
14	1159.5	28.7	1876	2	AAV12197	Human neu
15	1159.5	28.7	1876	6	ABS54875	Neuronal
16	1159.5	28.7	1876	6	ABS54875	Human neu
17	1159.5	28.7	1876	8	ABV73248	Human neu
18	1155.5	28.6	1590	2	AAV44687	Human neu
19	1153.5	28.5	1509	4	AAC90385	V274T var
20	1149.5	28.4	1509	4	AAC90385	Mutant hu
21	1143.5	28.3	1509	4	AAC90386	Mutant hu
22	1129	27.9	1964	6	ABZ11298	Mutant hu
23	1015.5	25.1	5096	4	ABL13732	Human pol
24	973	24.1	1915	4	AAD20962	Caenorhab
25	935	23.1	1416	4	AAC90382	Chimeric
26	848.5	21.0	2277	2	AAV12199	Human neu
27	848.5	21.0	2277	2	AAQ90387	Human neu
28	848.5	21.0	2277	6	ABS54870	Alpha 2 s
29	848.5	21.0	2277	6	ABV73243	Human neu
30	848.5	21.0	2664	2	AAT48235	Human neu
31	848.5	21.0	2664	6	ABK92165	Neuronal
32	848.5	21.0	2664	8	ADA10854	Prostate
33	848.5	21.0	2664	8	ADA10854	Human neu
34	848.5	21.0	2666	9	ADC71170	Human 205
35	847.5	21.0	2319	4	ABL04391	Human 205
36	845	20.9	1908	2	AAT48236	Neuronal
37	845	20.9	1908	8	ADA10856	Human neu
38	844.5	20.9	2210	4	ABL10275	Human neu
39	841.5	20.8	14668	4	ABL07798	Drosophil
40	841	20.8	2398	4	ABL26491	Drosophil
41	840	20.8	1584	7	ADA83809	Drosophil
42	834	20.6	1896	6	AAL45867	Human CHR
43	831.5	20.6	2385	4	ABL11821	Modified
44	829.5	20.5	1503	4	AAD20961	Drosophil
45	826.5	20.4	2170	4	ABL09701	Caenorhab

ALIGNMENTS

RESULT 1

Aaz24475

ID Aaz24475 standard; cDNA to mRNA; 2886 BP.

XX AC Aaz24475;

XX AC Aaz24475;

DT 17-FEB-2000 (first entry)

XX DE D. melanogaster acetyl-choline receptor DNA from clone Da7.

XX KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;

XX KW neurotransmission; plant protection agent; conductance; ACHR; ds.

XX OS Drosophila melanogaster.

XX FH Key Location/Qualifiers

XX CDS 372..2684

XX FT /*tag= a

XX FT /product= "acetyl choline receptor."

XX PN DE19819829-Al.

XX XX 11-NOV-1999.

XX PF 04-MAY-1998; 98DE-01019829.

XX PR 04-MAY-1998; 98DE-01019829.

XX PA (FARB) BAYER AG.

XX PI Adamczewski M, Oellers N, Schulte T;


```
QY 495 AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIle 514
Db 983 AACGAGATCTACTAACTGTGTCCGAGGACCATACATCGATCAACGTTTCGGTGTG 1042
QY 515 IleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla 534
Db 1043 ATCCGAGGAAACGCTACTACTCTTCAATCTGATCGTCCCTGCTGCTCATGCC 1102
QY 535 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLeuSerLeuGly 554
Db 1103 TCCATGGCTCTATTGGGTTTCACCTTGCTCCAGACTCCGAGAAAGTTGCTTTAGT 1162
QY 555 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 574
Db 1163 GTGACGATATTACTGTGCTGACGTTTCTCAACATGGTGGGAGACCATGCCAGC 1222
QY 575 ThrSerAspAlaValProLeu----- 581
Db 1223 ACGTCGAGCGCGTGCCCTTGCTCGGCACCTACTTCAACTGCATCATGTTTCATGGTGT 1282
QY 581----- 581
Db 1283 TCCTCCGCTGCTCCACCATACTGATCCTCAACTACCACCGGCGACGACACTCAC 1342
QY 582----- 582
Db 1343 GAAATGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1402
QY 598 SerArgProGly-----ArgProLeuIleLeuGluPheThrThrProCysSer 614
Db 1403 TCACGCGCGCTCGGCGACGACGCGCGCGCGCGCGCTACCTCCGCGCGC----- 1456
QY 615 AspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSer 634
Db 1457----- 1457
QY 635 SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArg 654
Db 1481 TCCAGTGTCTCTAGCAACGCTGCTGCATCTGATGACGACTTCGCCAC----- 1531
QY 655 ProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly 674
Db 1532----- 1532
QY 675 GlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThr 694
Db 1571 GGGGGTGAGGAGATGGCGCGGG----- 1600
QY 695 HisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuLeuLysGluIle 714
Db 1601 GGCACAGTTGTTTC-----GGTGTGACTACGAGCTCTCCCTCATCTGAGGAGATT 1654
QY 715 ArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrp 734
Db 1655 AGAGTCATCACAGATCAGATGCGCAAGGACGACGAGATGCGGACATTTTCGCGCACTGG 1714
QY 735 LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIle 754
Db 1715 AAGTTGCGCGCATGCTGTGACGACACTGTGCTTATTATCTTTTACCTGTTTCAATC 1774
QY 755 LeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIleValSer 770
Db 1775 ATCGCCACGCTAGCGGTGCTGTGTCGCGCCACACATCATGCTGTGTCG 1822
```

RESULT 3

AAZ24477
ID AAZ24477 standard; cDNA to mRNA; 3109 BP.

AC AAZ24477;

XX 17-FEB-2000 (first entry)

XX H. virescens acetyl-choline receptor DNA from clone Hva7-2.

```
XX  
KW  
XX  
XX  
OS Heliothis virescens.  
XX  
XX DE19819829-A1.  
XX  
PD 11-NOV-1999.  
XX  
XX 04-MAY-1998; 98DE-01019829.  
XX  
XX 04-MAY-1998; 98DE-01019829.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Adamczewski M, Oellers N, Schulte T;  
XX  
XX WPI; 2000-014207/02.  
XX  
XX P-PSDB; AAY50816.  
XX  
XX New nucleic acid encoding a nicotinic acetylcholine receptor from  
XX insects, used to identify potential insecticides.  
XX  
XX Claim 1a; Page 19-22; 26pp; German.  
XX  
XX This invention describes a novel nucleic acid (NA) encoding a nicotinic  
XX acetyl-choline receptor (I) from insects which can be used as an  
XX insecticide. Inhibitors of (I) interfere with neurotransmission. (I)  
XX (also vectors containing it, its regulatory regions, and antibodies  
XX directed against (I)-encoded proteins) are used to screen for: (a) plant  
XX protection agents that alter conductance of ACHR, potentially useful as  
XX insecticides, or (b) genes which encode polypeptides that are involved in  
XX formation of functionally related ACHR in insects. (I) are also used to  
XX isolate and characterize the specified regulatory regions and for  
XX recombinant production of (II). This sequence encodes an acetyl-choline  
XX receptor isolated from Heliothus virescens  
XX  
SQ Sequence 3109 BP; 923 A; 610 C; 742 G; 834 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.: 4,74e-135 Length: 3109
Score: 1609.00 Matches: 319
Percent Similarity: 71.62% Conservative: 57
Best Local Similarity: 60.76% Mismatches: 65
Query Match: 39.80% Indels: 84
DB: 3 Gaps: 10

US-09-303-232-2 (1-770) x AAZ24477 (1-3109)

```
QY 295 LeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArg 314
Db 122 CTGCTGGCTTTCCTGCCCGTATCCGAGCAA-----GGTCTCACGAGAAGA 169
QY 315 LeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSer 334
Db 170 CTCCTGAACGCGTGTCTGGCGAATCAACACCTGGAGCGACCGTGCCCAACGAGAGC 229
QY 335 AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGluLys 354
Db 230 GAACCGCTAGAGGTGAGGTTCGGCTTGACCTTTCGACCAATCATTTGACGTGACGAGAG 289
QY 355 AsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArg 374
Db 290 ATCAACTACTTATAACCAATATATGCTGTGCTGAGTGGAGTGAATGACTACAACTGAGG 349
QY 375 TrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProHisArgIleTrp 394
Db 350 TGAACGACGACGAGTATGCGCGGTCAAGGACCTCAGGATCAGCCCAACAAAGTTGTGG 409
QY 395 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr 414
Db 410 AAGCCGAGCGTCTCTTATGTATATAGTGTGCTGACGAGGGTTTGACGGGACCTACGAGAC 469
```

```
QY 415 AsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIlePheLysSer 434
Db 470 AACGTGGTGTGAGAGCGGCGAGTTCCTGTAGTCCACCTGGCATATTCAGAGC 529
QY 435 ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe 454
Db 530 ACATGCAAGATGACATCGCGTGGTTCCCTTCGACGACCAACACTGTGATATGAGTTC 589
QY 455 GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGly 474
Db 590 GGTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAAGATGAGGACGGCGC 649
QY 475 AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuValProGlyLysArg 494
Db 650 GATCTATCGGACTTCATACAAATGGGAGTGGTATCTAATAGGAATCCAGGCAAAAG 709
QY 495 AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIle 514
Db 710 AACACATAACATAGCGTGTGCTGCCGAGCCCTAGGTGGAGCTCACCTTCACCATCATG 769
QY 515 IleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla 534
Db 770 ATAGAAGACGAACCTTGTACTACTTCTCAACTGATCGTCCCGTGGTGTGATCTCA 829
QY 535 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly 554
Db 830 TCGATGGCACCCTCGGCTTCACATCGCACACAGACTCCGGAGAGAACTCACACTTGGG 889
QY 555 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 574
Db 890 GTCACTATCTCTATCGCTGACGGTGTCTCAACCTGGTAGCGGACCCCTGCCACAG 949
QY 575 ThrSerAspAlaValProLeu----- 581
Db 950 GTCTCCGACGCTATCCCCCTGTTAGGACGTACTTCAATTGCATCATGTTTCATGGTAGCG 1009
QY 581 ----- 581
Db 1010 TCGTCTGTGTACTGACTGTGGTGTACTCAATTACCCACCATCGAACAGCTGATATCAT 1069
QY 582 -----TrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMet 597
Db 1070 GAAATGCCACAGTGGATAAATCAGTATTCCTACATGGTTGCCATGGATATCGGATG 1129
QY 598 SerArgProGlyArgProLeuLeuLeuGluPheProThrThrProCysSerAspThrSer 617
Db 1130 TCGAGGCCGAGGAGAGATCACCGAAG-----ACTATAATCATGACACACGAGG 1180
QY 618 SerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSer 637
Db 1181 -----ATGAGGAGCTGGAATCGAATGAGGAGAGAGTGTGCGAATGCC 1219
QY 638 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThr 657
Db 1220 TTGCTGGCGAATGTTCTAGATATGATGATGACTTCAGACACGCGC-----CCTCCGCT 1273
QY 658 ProGly-----GlyThrLeuProHisAsnProAlaPheTyrArgThrVal 672
Db 1274 CCTAACAGTACTGCTCGACCGGGAATTTGGGACCTGGGTGCTCAATATTCGCGACCGAT 1333
QY 673 TyrGlyGlnGly-----AspAspGlySerIleGlyProIle 684
Db 1334 TTCGTGGTGTGCTGCTCGCTCCGTCACATGAGAGACGTGGC-----GGCGGCTG 1387
QY 685 GlySerThrArgMetProAspAlaValThrHisThrCysIleLysSerSerThrGlu 704
Db 1388 GGTAGC-----CACCATCGC----- 1402
QY 705 TyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAsp 724
Db 1403 ---GAGCTGCACCTACTACTGAGAGAGCTGCTGATCATCACGCGCCAGGATGAAGAGCT 1459
```

```
QY 725 AspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArgLeu 744
Db 1460 GATGAGGAAGCGGAGCTGATCAGCTGCGAGTTCGCTGCGATGGTTCGTGATAGTTT 1519
QY 745 CysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAla 764
Db 1520 TCCCTGTTTCGTTTCACACTTTTCAATCATCGCGACAGTAGCTGTCTCTGTTATCGGCA 1579
QY 765 ProHisIleIleVal 769
Db 1580 CCGCATATCATCGTG 1594
RESULT 4
ABU27131
ID ABL27131 standard; DNA; 984 BP.
XX
AC ABL27131;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32866.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WIPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 32866; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL6175) and the encoded proteins (ABBS7737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 984 BP; 332 A; 268 C; 225 G; 159 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8.57e-126 Length: 984
Score: 1498.50 Matches: 297
Percent Similarity: 98.69% Conservative: 4
Best Local Similarity: 97.38% Mismatches: 3
Query Match: 37.06% Indels: 1
DB: 4 Gaps: 1
US-09-303-232-2 (1-770) x ABL27131 (1-984)
QY 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
```

Mon May 10 12:05:14 2004

us-09-303-232-2.rng

Db 1 ATGAATAATGCACAACTGAACTGACTGAAAGTTGACGATGATGAGCTGTGGCTGGCAGTA 60
Qy 21 ArgLeuAlaHisCysSerSerAsnPhSerSerSerSerSerThrArgThrThrSerSer 40
Db 61 AGATTAGCGCACTGACGAGCAACATCAGCAGCAGTAGCAGCAACAGAACACACGAGCAGC 120
Qy 41 AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
Db 121 AACAGAGGGCAACACCGAGCACTCACAACACTGCAACCAAGAGAGCTTAAGTACAAACAC 180
Qy 61 HisSerAsnLeuAlaSerGlnGlnHisAsnSerGlnGlnGlnGlnProAlaSerLysAsp 80
Db 181 CACAGCAACATTTGCAAGCGAGCAGCAACATAGCCAGCAACAGGAGCCAGCATCGAAGGAC 240
Qy 81 GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp 100
Db 241 GAGGATGTAGCAACACCGTAGAGCAATCACCAGCAGCAGCGCATCTGCAACAGCTAGAC 300
Qy 101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla 120
Db 301 AGCAGCAACATGTTGTGCCAAAGACAGCCGAGCAGCAACTGCTGCCGCGGATGAAGCA 360
Qy 121 ThrThrGlnGlnProThrAsnLeuArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
Db 361 ACAACCCCAACCAACCAACATAGACTGTGTGACGCAAGCGACCAACGATTGGCTGCGC 420
Qy 141 ArgArgLysArgLysProAlaThrProAsnGlnThrAspLysLysGlnGlnGlnLeu 160
Db 421 CGACGAAAGAAACACCGACCAACCCCAACCAAGATATCAAGAAACACAGCAACTT 480
Qy 161 SerMetProProPheLysThrArgLysSerThrAspThrThrThrProAlaAlaThr 180
Db 481 AGCATGCTCTCTTCAAAACGGCGCAATCCACGGACACTACAGCACACAGCAACATA 540
Qy 181 ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGluPheSerLys 200
Db 541 ACCAGCTGTCCGACAGCCACCTACATGCAATGTCAGCGCAGCGACCAATGAGTTTCAGTAT 600
Qy 201 ProfileSerArgHisAspArgValSerThrAlaThrPheAlaTrpValLeuHisValLeu 220
Db 601 CCGATATCGAGACATGATAGATATCCACGGCCACATTCGCCCTGGGTGTGATGTGCTG 660
Qy 221 GlnValLeuLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu 240
Db 661 CAGGTGCTGCTGCTGCTGCTGCAACAGTGGCACTTCAGTGCACAGCGATCGGTGCTA 720
Qy 241 LeuPheArgArgLeuAlaAlaSerThrLeuAlaPheLysSerThrLeuGlySerPheAla 260
Db 721 CTGTTTCAGAAAGGATCGCAGCGAGCACCATCGCCTTCAATTCCTATTATTAGCGAGCTT 780
Qy 261 AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 279
Db 781 GCGCAACTCAGGAGCAGCAGCAGTAGTAGCAGCAGCAACAGCAGCAACAGCGCAACAGCAG 840
Qy 280 SerThrGlnLeuLeuAsnGlyLeuAsnLysHisSerTrpPheLeuLeuLysLeuLys 299
Db 841 AGCAGCGCAATATTAACCGACTTAATAAACACTCATGATATTTTATTGATATATTG 900
Qy 300 AsnLeuSerAlaLys 304
Db 901 AATTTATCTGCTAAA 915

RESULT 5
ABL07231
ID ABL07231 standard; cDNA; 1540 BP.
AC
AC ABL07231;
XX
XX ABL07231;
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 16175.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
EN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-65860/75.
DR P-PSDB; ABB63128.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 16175; 2lpp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1540 BP; 412 A; 390 C; 376 G; 362 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,33e-122 Length: 1540
Score: 1466.50 Matches: 302
Percent Similarity: 65.09% Conservative: 41
Best Local Similarity: 57.31% Mismatches: 73
Query Match: 36.27% Indels: 112
DB: 4 Gaps: 8
US-09-303-232-2 (1-770) x ABL07231 (1-1540)
Qy 291 SerTrpPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 310
Db 71 TCGCTGTTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127
Qy 311 HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal 330
Db 128 CATGAAAAGCGCTGCTGAACCATCTGCTGTCCACTACATACGCTGAGGAGCAGCCGTG 187
Qy 331 LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnLeuLeuAsp 350
Db 188 GCCAATGATCGAGCCCTCGAGGTTAAGTTCCGACTGACGCTGCGAGCAGATCATCGAC 247
Qy 351 ValAspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp 370
Db 248 GTGGATGAAAAGAAATCAGATTCTGACCAAAATGCGTGGTTAAATTTGAGGTGGAACGAC 307
Qy 371 MetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgLysProPro 390
Db 308 TACAATCTCGCTGGAATGAACGGAATACGCGGGGTCAAGGATCTACGATCACGCC 367
Qy 391 HisArgLysTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410
Db 368 AACACGCTGTGGAAGCCCGACGCTGCTCATGTATCAACAGCGCGGATGAGGATTCGATGCG 427

QY 411 ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGly 430
Db 428 ACGTATCACCAACATTGTGTCTCAACATACAGCAGTTGTCTGTACGTGCCCTGGT 487
QY 431 IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspGlnArgCys 450
Db 488 ATCTTTCAAGAGACATGCAAGATAGACATCACGTGTTCCCATTTGATGATGACCAATTGC 547
QY 451 GluMetLysPheGlySerThrTrpTyrAspGlyPheGlnLeuAspGlnLeuGlnAsp 470
Db 548 GAAATGAATTCGGTAGTTGGACTTACCATGGAATTCAGTTGGATTTGGTTGAATTCC 607
QY 471 GluThrGlyAspIleSerThrValLeuAsnGlyGluTrpGluLeuGly--- 489
Db 608 GAAGATGAGGGGATCTTCCGATTTCATAACAAATGGCGAGTGTACTTGTGTTAC 667
QY 490 -ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle 509
Db 668 CATGCCGGAAGAAGATAGATAGTCTACGCTGTGCCAGAACCATATGTCGATAT 727
QY 509 eThrPheAlaIleIleIleArgArgThrLeuTyrTyrPheAsnLeuIleIlePr 529
Db 728 CACCTTACTATACAAATCTGCGCGTACATATATTATTTTCAATTTAATTTGTCGC 787
QY 529 oCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyG1 549
Db 788 ATGTGTGCTAATCTCATGATGGCCTACTGGCTTTCATTTGCGCGGATTCGGCGA 847
QY 549 uLysLeuSerLeuGlyValThrIleLeuSerLeuThrValPheLeuAsnMetValAl 569
Db 848 GAAACTGACGTGGCGTAACTACTACTATCATTAACAGTATTTCTAAACCTTGTGCG 907
QY 569 aGluThrMetProAlaThrSerAspAlaValProLeu----- 581
Db 908 CGAGTCCATGCCGACACGTCGAGTGTGTTCTCT-TATAGGTACACACAGATATT 966
QY 581 ----- 581
Db 967 AAAACGCTATAGCTTAAAGTACGACCGATCGGTGGACTATCGACCACTACTACAGTG 1026
QY 581 ----- 581
Db 1027 TCGAGCAGAAAGCACCTACTTCAATTGCATCATGTTTCATGTCGCTCGTGGTGTCT 1086
QY 582 -----Tr 582
Db 1087 GACAGTAGTGTGCTCACTACCACTCCGACAGCGGACATTCACGAGATGCCACCGTG 1146
QY 582 pIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMetSerArgProGlyAr 602
Db 1147 GATCAAGTCGGTTTCTCTACAAATGGCTGCCCTGGATCTTGGATGGTGCACCGGTG 1206
QY 602 qProLeuIleLeuGluPheProThrThrProCysSerAspThrSerSerGluArgLysH1 622
Db 1207 CAGATT-----ACACGCAAAAC 1224
QY 622 sGlnIleLeuSer-----AspValGluLeuLysGluArgSerSerLysSerLe 638
Db 1225 AATATTATTAAAGCAATCGCATGAGGAGCTGGAGCTAAAGAGCGCTCTCCCAATCCCT 1284
QY 638 uLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThrPr 658
Db 1285 GCTGGCAATGTCTCGACATCGACAGCAGCTTCCGCGCACACATA----- 1330
QY 658 oGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGlyGlnGlyAspAs 678
Db 1331 -----TCTGGCTCCCA 1341
QY 678 pGlySerIleGlyProIleGlySerThrArgMetProAspAlaVal-----ThrHisIstH 697
Db 1342 AACCGCATTTGGCTCGCGCCAGCTTCGGTCGGCCCAACCGGTGGAGGAGATCACAC 1401
QY 697 rCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheI1 717

Db 1402 GCCATC---GGCTCAATCAAAAGATCTTCATTAATCTCAAGAATTCGAATTAT 1458
QY 717 eThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAl 737
Db 1459 TACGGCGGGATGCGCAAGAGCTGACGACGAGCGAATGATCGCGGATTTGAAGTTGCG 1518
QY 737 aAlaMetValValAspArg 743
Db 1519 GGCATGTTGTGGATAGG 1537
RESULT 6
ABL27130
ID ABL27130 standard; DNA; 3144 BP.
XX
AC ABL27130;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32863.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-658660/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 32863; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3144 BP; 1056 A; 676 C; 680 G; 732 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3,12e-121 Length: 3144
Score: 1456.50 Matches: 300
Percent Similarity: 86.89% Conservative: 5
Best Local Similarity: 85.47% Mismatches: 12
Query Match: 36.03% Indels: 35
DB: 4 Gaps: 3
US-09-303-232-2 (1-770) x ABL27130 (1-3144)
QY 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
Db 1001 ATGAAATATGCACACTGAACCTGACTGAAGTTGACGATGATGCTGGCTGGCAGTA 1060


```
Db 295 AATCTTCGATGGAATTCGAGTCAGTTCGGTGGTGGCGGATCTCGGAATTCGCGCCACAT 354
Qy 392 ArgileTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db 355 CGCCTATGGAACCGGATGTACTGATGTACAAACAGTCGCGCAGGGCTTCGATGGAACG 414
Qy 412 TyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIle 431
Db 415 TACGCCAATATGTTGGTTCGCNATATGGAGCTGTCTGATCGTACCCGCCAGGTATA 474
Qy 432 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 451
Db 475 TTAAAGTCAACGCTGAAGATCGACATTACGTGGTTCATTCCAGATCGAGATGTGAA 534
Qy 452 MetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGlu 471
Db 535 ATGAAATTTGGTTCGTGGACCTACGATGGGTTCAGTTGGACCTGCAGTTGCAGACGAA 594
Qy 472 ThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValPro 491
Db 595 GCTGGTGGCGACATTTCTAGCTTTATACCAATGCGGAATGGGACTTGTAGGTGTGCC 654
Qy 492 GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
Db 655 GGTAAACGAAATGAATCTACTATAATTTGCTGCCAGAACCTTATATTGACATACATTC 714
Qy 512 AlaIleIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysVal 531
Db 715 GCCATTTTGTATGAGCGCGAAACGTTGTACTATTATTTTCAATCTGATGTGCCGTG 774
Qy 532 LeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeu 551
Db 775 CTGATCGCTCCATGGGCACCTGCTAGGGTTTACATGCCACCAGATTCCTGGTGAAGCTT 834
Qy 552 SerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThr 571
Db 835 TCGCTTGAGTTACAAATCTATTATTCGTTACAGTCTTCCTCAACATGGTGGCCGAAACA 894
Qy 572 MetProAlaThrSerAspAlaValProLeu 581
Db 895 ATGCCGCGGACCTCCGATGCGGTACCGGTG 924
```

RESULT 8

ABL07799
ID ABL07799 standard; cDNA; 803 BP.

AC ABL07799;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 17879.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX W0200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB63696.

XX
PT
PT
PT
PS
XX
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ
Alignment Scores:
Pred. NO.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-09-303-232-2 (1-770) x ABL07799 (1-803)
Qy 352 AspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMet 371
Db 1 GACGAGAAAATCAATTCGTCAGTCACTAATGTGGTTAAACCTGAGTGGACGACATG 60
Qy 372 AsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProHis 391
Db 61 AATCTCGCTGGACACCTCGACTATGGCGAGTTAAGGATCTGCGAATACCCCGCAT 120
Qy 392 ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db 121 CGCATCTGGAAGCGGACGCTGATGTACACAGTCGGATGAGGATTTGATGGCACC 180
Qy 412 TyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIle 431
Db 181 TACCAGACGAACGTTGGTGGTGGCGAACACACGGCTCGTGTCTATACGTTCCCGCGGATC 240
Qy 432 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 451
Db 241 TTCAAGTCGAGCTGCAAGATCGACATCACGTGGTTCCTTCGATGACCGGTCGAG 300
Qy 452 MetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGlu 471
Db 301 ATGAAGTTCGGCAGTTGGACCTACGACGATTCAGGTT----- 339
Qy 472 ThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValPro 491
Db 340 -----TGG-----TTCAAGTGTGCC 354
Qy 492 GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
Db 355 GGCBAACGTPAACGAATCTATTACACTGTCGCGGAACCTATATAGACATCACCTTC 414
Qy 512 AlaIleIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysVal 531
Db 415 GCCATCATCATCCCGCAGCAACACTGTACTATTTCTTCAACCTGATCATACCTTGTGTA 474
Qy 532 LeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeu 551
Db 475 CTGATTTGCCCTCCATGGCCTTGGCTCGGATTCACCTCGCCGACAGATTCGGGTGAAAAATTA 534
Qy 552 SerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThr 571
Db 535 TCACTGGGTGTACCAATCTGCTCTCGTACCGGTGTTCTGAATATGTTGTCGAGACA 594

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
Claim 1; SEQ ID NO 17879; 21pp + Sequence Listing; English.
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU16176-ABU16175), expressed DNA sequences (ABU16176-ABU16175) and the encoded proteins (ABBS7737-ABBS7742). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Sequence 803 BP; 211 A; 199 C; 190 G; 203 T; 0 U; 0 Other;

```

Qy 572 MetProAlaThrSerAspAlaValProLeuTrpIleArgIleValPheLeuCysTrpLeu 591
Db 595 ATGCGGCTACTTCGATGCGTGCCATG-----624
Qy 592 ProTrpIleLeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrThr 611
Db 624 -----624
Qy 612 ProCysSerAspThrSerSerGluArgGlyHisGlnIleLeuSerAspValGluLeuLys 631
Db 625 -----CTGGAACGTGACATAATCAATTTAATCAAAATCAAA-----660
Qy 632 GluArgSerSerLysSerLeuAlaAsnValLeuAspIleAspAspPheArgHis 651
Db 660 -----660
Qy 652 AsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThr 671
Db 660 -----660
Qy 672 ValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAsp 691
Db 660 -----660
Qy 692 AlaValThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeu 711
Db 660 -----660
Qy 712 LysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAla 731
Db 661 -----CAGCTACGTAAAGATGACGAGTGCAATGACATTGCC 696
Qy 732 AsnAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMet 751
Db 697 AATGATTGCAATTTGCGAGCTATGCTGTGACAGCTGTGCTTATCATATTCACAATG 756
Qy 752 PheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHis 766
Db 757 TTCACAATATTAGCCACATAGCTGTACTATCAGACCACAT 801

RESULT 9
ID AAT59196 standard; cDNA; 2769 BP.
AC AAT59196;
XX
DT 17-JUN-1997 (first entry)
DE Neuronal alpha-bungarotoxin binding protein alpha subunit cDNA.
KW Neuronal alpha-bungarotoxin binding protein alpha 1; cholinergic;
KW ligand binding; ion channel; ss.
XX
OS Gallus sp.
XX
FH Key Location/Qualifiers
FT CDS 71..1513
FT /tag= a
FT sig_peptide 71..136
FT /tag= b
FT mat_peptide 137..1510
FT /tag= c
XX
PN US5599709-A.
XX
PD 04-FEB-1997.
XX
PF 28-SEP-1989; 89US-00413947.
XX
PR 28-SEP-1989; 89US-00413947.
XX
PA (SALK ) SALK INST BIOLOGICAL STUDIES.

```

```

XX Lindstrom JM, Schoepfer RD;
XX WPI; 1997-118297/11.
XX P-PSDB; AAM12368.
XX New isolated neuronal alpha-bungarotoxin-binding protein DNA - used to
XX screen cholinergic agents and other drugs which may affect ligand
XX binding, ion channel or other activities of the protein.
XX Claim 1; Fig 2A-B; 18pp; English.
XX
XX 2 cDNA clones (AAT59196 and AAT59197) respectively code for the alpha
XX subunit (AAM12368) and alpha2 subunit (AAM12369) of chick neuronal alpha-
XX bungarotoxin binding protein (ABBP). They were isolated from an 18-day
XX embryo chick brain cDNA library using a probe (see also AAT59198) based
XX on the N-terminal amino acid sequence of chicken brain ABBP. The probe
XX isolated partial clone pCh29-1, which encoded the N-terminal portion of
XX alpha1. A subclone, pCh29-3 (ATCC 40641), was used to rescreen the
XX library, yielding clone pCh31-1 (ATCC 40640), which encoded the entire
XX alpha2 sequence. A probe based on the C-terminal region of pCh31-1 was
XX used to obtain clone pCh34-1 (ATCC 40639), encoding the C-terminal portion
XX of alpha1. The cDNA clones can be used as probes to identify further ABBP
XX subunits, and in the recombinant prodn. of ABBP
XX
SQ Sequence 2769 BP; 790 A; 604 C; 571 G; 804 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.16e-96 Length: 2769
Score: 1179.50 Matches: 237
Percent Similarity: 61.08% Conservatives: 80
Best Local Similarity: 45.66% Mismatches: 127
Query Match: 29.17% Indels: 75
DB: 2 Gaps: 9

US-09-303-232-2 (1-770) x AAT59196 (1-2769)
Qy 295 LeuLeuIleTyrLeuAsnLeuSerAla-----LysValCysLeuAlaGlyTyrHis 311
Db 86 CTGATGCTGTGCTGCTGCGCGCGGGCGGCTGCGGAGTCCCTCGAAGGAGATTC 145
Qy 312 GluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeu 331
Db 146 CAAAGGAAGCTGTACAGGAGCTGCTGAAGAACTACAACCTCTCGAAGCAGCAGTGTGA 205
Qy 332 AsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspVal 351
Db 206 AATGACTCCAGCGCTCAGTCTATTCTACTCTCAGCCTCATGCGATCATGATGTG 265
Qy 352 AspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMet 371
Db 266 GATGAAAAGAAATCAAGTATTAAACAACAACATCTGGCTACAAATGTCTGGACAGATCAT 325
Qy 372 AsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProHis 391
Db 326 TACTTACAGTGAATGTGCTGCAATACCTCGAGTGAAGACGTCGTTTCTCTGATGA 385
Qy 392 ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db 386 CTGATTGGAGCCAGATATTCTTCTCTATACAGTGTGATGAAGATTGTGATGTACA 445
Qy 412 TyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIle 431
Db 446 TTTCACACTAATGTTTGTAGTCAATTTCTTCGGGACACTGCCAATATCTGCCACGACATA 505
Qy 432 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGluArgCysGlu 451
Db 506 TTTAAAGCTCATGCTACATAGCGTGGTGTTCATTTCATTGATGTTGAGAGTGCAT 565
Qy 452 MetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGlu 471
Db 566 CTGAAGTTTGGATCTTGACATATGAGGCTGCTCTTAGACTTACAAATGCAAGAA--- 622

```



```
Db 122 TCCTTCTTT-----AAAGTATCTCAGCAGGCTGAAGCCACAGAGCGCTTG 166
Qy 316 LeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSerAsp 335
Db 167 TACAGAGATCTGCTGAGAAACTACAATCGTCTGGAGCGCCAGTCATGAATGACTCCAG 226
Qy 336 ProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGluLeuAsn 355
Db 227 CCCATTGAGTTGAGCTTCAGCTTCTTGTCTGCAGATAATTGATGGATGAGAGAAT 286
Qy 356 GlnLeuLeuValThrAsnValTyrLeuLeuLeuTyrAsnAspMetAsnLeuArgTyr 375
Db 287 CAGGTGTTGATTACAATGCTTGGCTGCAGATGACTGGGTGATATTACTTCTCTGG 346
Qy 376 AsnThrSerAspTyrGlyValLeuAspLeuArgIleProProHisArgIleTyrIys 395
Db 347 GATCAGTACGAATAACCCGGGTGTGCAGAACTTGGGATTTCCATCCGACCAAGATTGGGTA 406
Qy 396 ProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsn 415
Db 407 CCTGATATCTTCTATAACAGTGGGATGAAAGATTCGATGCAACATTTTCACACAAT 466
Qy 416 ValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLeuSerThr 435
Db 467 GTGCTGGTGAATTAATCTCGATCCTGCAATATATTCCTCCAGGCAATTTTCAAGAGCACA 526
Qy 436 CysIysIleAspIleThrPheProPheAspAspGlnArgCysGluMetLysPheGly 455
Db 527 TGTATCATTTGATGTCGCGTGGTCTTCCCTTTGATGTCAGAAAGTGTGATTTGAAGTTTGGC 586
Qy 456 SerTyrThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGlyAsp 475
Db 587 TCTTGGACTCAGTGGTGGCTGATTCACCTGCAGATGCTTGAG-----GCTGAT 637
Qy 476 IleSerSerTyrValLeuAsnGlyGluTyrProGlyLeuLeuGlyValProGlyLysArgAsn 495
Db 638 ATTTCCAACTACATCTCAATGGAGAGTGGGATTTAGTGGGTGTCCTCCAGGAAAGAGGAAT 697
Qy 496 GluIleTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIleIle 515
Db 698 GAGTTGTACTATGAATGCTGTAAAGAACCATATCCAGATGTCACATACACCATCACCATT 757
Qy 516 ArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAlaSer 535
Db 758 CGACAGCACTCTCTACTATGCTTGAACCTACTGATTCCTGTGTTCTCATATCTGCG 817
Qy 536 MetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyVal 555
Db 818 TTGGCACTGCTGTATTCTTTGCTGCTGATTCCAGGAGAGAAGATTCTTTAGGTATC 877
Qy 556 ThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThr 575
Db 878 ACTGTTCTGCTTTCCTCACTGATTCATGCTGTTGTGGCTGAGATCATGCTGCAACT 937
Qy 576 SerAspAlaValProLeu----- 581
Db 938 TCTGATTCAGTCCCACTAATAGCTCAGTATTTTGTGATCATCATGCTATTGTTGGTCTG 997
Qy 581 ----- 581
Db 998 TCTGTGTCGTAAACGGTCTGCTTCTGCAGTTTACCACCATGACCCACAGCAGGAAG 1057
Qy 582 -----TyrIleArgIleValPheLeuCysTyrProTyrProTyrIleLeuArgMetSer 598
Db 1058 ATGCCACAGATGGGTCGGTCTGCTGATTCGTGATTCGTGCTGTTTACGTATGAAA 1117
Qy 599 ArgProGly-----ArgProLeuIleLeuGluPheProThrThrProCysSerAsp 615
Db 1118 AAACCTGGGAAATATAAGCCCTCTCTTGGCAATAT----- 1156
Qy 616 ThrSerSerGluArgLysHisGlnIleLeuSerValGluLeuLysGluArgSerSer 635
```

```
Db 1157 ---AGCTATCCCAACACACCATCCAAAGCTGAAAAACACAGAGATG----- 1198
Qy 636 LysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgPro 655
Db 1199 -----AATGTCCTA----- 1207
Qy 656 MetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly--- 674
Db 1208 -----CTGGG-----CACCAGCCCAAGCAATGCAACATGATTTATAGCTAC 1249
Qy 675 -----GlnGlyAspAsp-----GlySerIleGlyPro 683
Db 1250 CACACAATGAAATATCCATGCTGCCCCAGACAATGATCTGGGCAGCAGAGAGTGGAAAG 1309
Qy 684 IleGlySerThrArgMetProAspAlaValThrHisIleThrCysIleLeuSer----- 701
Db 1310 ATTACTTGGCCCTTGTTCAGAGAT-----AACAGCATGTTCAAAAAAAGCTTTAATG 1363
Qy 702 SerThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGlnLeu 721
Db 1364 GATACCATCCAGTGATTTGTGAAGATACTGGAGGAAGTTCAGTTTCATAGCAATCGCTTC 1423
Qy 722 ArgLysAspAspGluCysAsnAspIleAlaAsnAspTyrPheAlaMetValVal 741
Db 1424 AGGAAGCAAGATGAGGTGAGAGATCTGCAGTGAAGAGTTTGACGCTGCTGCATA 1483
Qy 742 AspArgLeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeu 761
Db 1484 GACAGATTATGCTGTTGCTGTTTACCTTTTTCCTATTCATTCACACATTTACATACTC 1543
Qy 762 LeuSerAlaProHisIleIle 768
Db 1544 ATGCTGCTCCCAACTTTATA 1564

RESULT 11
AAC58395
ID AAC58395 standard; cDNA; 1509 BP.
XX AC AAC58395;
XX XX 29-JAN-2001 (first entry)
DE Human PRO2145 nucleotide sequence SEQ ID NO:76.
KW Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
KW proliferation; tumorigenesis; identification; cancer; cytostatic;
KW neotropic; neuroprotective; antiinflammatory; immunosuppressive;
KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;
KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;
KW hypothalamic disorder; glandular disorder; macrophagal disorder;
KW epithelial disorder; stromal disorder; blastocellic disorder;
KW inflammatory disorder; immunologic disorder; ss.
XX Homo sapiens.
OS WO200053755-A2.
XX XX
XX XX 14-SEP-2000.
XX XX
XX XX 06-JAN-2000; 2000WO-US000376.
XX XX
XX XX 08-MAR-1999; 99WO-US005028.
XX XX 02-JUN-1999; 99WO-US012252.
XX XX 23-JUN-1999; 99US-0141037P.
XX XX 07-JUL-1999; 99US-0143048P.
XX XX 26-JUL-1999; 99US-0145698P.
XX XX 30-NOV-1999; 99WO-US028313.
XX XX 20-DEC-1999; 99WO-US030911.
XX XX 05-JAN-2000; 2000WO-US000219.
XX XX
XX (GETH ) GENENTECH INC.
XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
PI
```


Qy 744 LeuCysLeuIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSer 763
Db 1408 CTGTCCTCATGGCTTCTCGTCTTACCATCATCTGCACCATCGGCATCTGATGCG 1467
Qy 764 AlaProHisIleIle 768
Db 1468 GCTCCCAACTTCGTG 1482

RESULT 12
ID AAC90380 standard; cDNA; 1509 BP.
XX AC
XX AAC90380;
XX DT 14-MAR-2001 (first entry)
XX Wild-type human alpha7 ligand gated ion channel coding sequence.
XX Human; alpha7 nicotinic acetylcholine gated ion channel;
KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.
XX OS Homo sapiens.
XX WO200073431-A2.
XX 07-DEC-2000.
XX 25-MAY-2000; 2000WO-US011862.
XX 27-MAY-1999; 99US-0136174P.
XX (PHAA) PHARMACIA & UPJOHN CO.
XX Groppi VE, Wolfe ML, Berkenpas MB;
XX WPI; 2001-061524/07.
XX P-PSDB; AAB50012.

XX Special cell culture medium for treating cells and for inducing mammalian
PT cell lines to conduct calcium ions, comprising specified concentrations
PT of ions of sodium, calcium and potassium at specified pH.
XX Example 5; Page 60-61; 77pp; English.
XX The present sequence is the coding sequence for wild-type human alpha7
CC nicotinic acetylcholine gated ion channel. The human alpha7 ion channel
CC was used in the construction of an alpha7/5-hydroxytryptamine (5-HT3)
CC chimeric ligand gated ion channel (see AAC90382 and AAB50014). The
CC alpha7/5-HT3 chimeric ion channel can be expressed by recombinant cells
CC in the present invention, resulting in preferential calcium ion
CC conductance by the cells
XX SQ Sequence 1509 BP; 297 A; 451 C; 429 G; 332 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.22e-95 Length: 1509
Score: 1159.50 Matches: 237
Percent Similarity: 60.76% Conservative: 82
Best Local Similarity: 45.14% Mismatches: 115
Query Match: 28.68% Indels: 91
DB: 4 Gaps: 12

US-09-303-232-2 (1-770) x AAC90380 (1-1509)
Qy 297 IleTyrLeuAsnLeuSerAlaLys-----ValCysLeuAlaGlyTyrHisGluLys 313
Db 22 GTCTGGCTGGCGCTGGCGGCTCGCTCTGCACGTGCTCCCTGCAAGCGAGTTCAGAGG 81
Qy 314 ArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGlu 333
Db 82 AAGCTTTACAGAGAGCTGGTCAAGAACTACAATCCCTTGGAGAGCCCGTGGCCCAATGAC 141
Qy 334 SerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGlu 353

Db 142 TCGCAACCACTCACCGTACTCTCTCCGTGAGCCTCTGCAGATCATGTGAGTGGAG 201
Qy 354 LysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeu 373
Db 202 AAGAACCAAGTTTAAACCAACACATTTGGTGTGAAATGTCTTGGACAGATCATATTATTA 261
Qy 374 ArgTrpAsnThrSerAspTyrGlyValValAspLeuArgIleProProHisArgIle 393
Db 262 CAGTGGAAATGTGTGAGAATATCCAGGGGTGAAGACTGTTCCTTTCCAGATGGCCAGATT 321
Qy 394 TrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTrpGln 413
Db 322 TGGAAACCAAGACATTTCTCTATAACAGTGTCTGATGAGCGCTTTCAGCCACATTCAC 381
Qy 414 ThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIlePheLys 433
Db 382 ACTAACGTGTGGTGAATTTCTTGGGCATTTGCCAGTACCTGCTCCAGGATATTCAG 441
Qy 434 SerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLys 453
Db 442 AGTTCCTGCTACATCGATGATGCTGTTTCCCTTTGATGTGAGCACTGCAAACTGAG 501
Qy 454 PheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGly 473
Db 502 TTTGGTCTCTGTCTTACGGAGGCTGCTCTTGGATCTGCAGATGCAGGAG----- 552
Qy 474 GlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuGlyValProGlyLys 493
Db 553 GCAGATATCATGTGGCTATATCCCAATGGAGATGGACCTAGTGGGAATCCCGGCAAG 612
Qy 494 ArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIle 513
Db 613 AGGAGTGAAGGTTCTATGATGCTGCTCAAGAGCCCTACCCCGATGTCACTTCACATG 672
Qy 514 IleIleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIle 533
Db 673 ACCATGCGCGCAGGACGCTCTACTATGGCTCAACCTGTGTGATCCCTGTGTCTCATC 732
Qy 534 AlasSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeu 553
Db 733 TCCGCTCTCGCCCTGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 792
Qy 554 GlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetPro 573
Db 793 GGGATACAGTCTTACTCTCTTACCGTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
Qy 574 AlaThrSerAspAlaValProLeu-----ArgProLeuIleLeuGluPheProThrThrPro 581
Db 853 GCAACATCCGATTCGGTACCATTGATAGCCAGTACTTGGCAGCAGCACCATGATCATG 912
Qy 581 ----- 581
Db 913 GGCCTCTCGTGTGTGACGGTGATCGTGTGACGTACCAACACCACAGCCCGGAG 972
Qy 582 -----TrpIleArgIleValPheLeuCysTrpLeuProTyrIleLeuArg 596
Db 973 GGCAAGATGCCAAGTGGACAGATCATCTTCTGAACTGGTGGTGGTGGTGGTGGTGGTGG 1032
Qy 597 MetSerArgProGly-----ArgProLeuIleLeuGluPheProThrThrPro 612
Db 1033 ATGAAGAGGCGCGGAGGAGCAAGGTGCGGCGC-----GCCTGCCAGCAACAAGCAGCGCGC 1089
Qy 613 CysSerAspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGlu 632
Db 1090 TGCAGCTCGCCAGTGTGGAG-----ATGAGCCCGTGGCGCGCGCGCGCGCG 1134
Qy 633 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHis--- 651
Db 1135 GCCAGCAACAGGAACTGCTGTAC-----ATCGGCTTCGCGGCGCGCTG 1176
Qy 652 -----AsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPhe 668

Db 1177 GAGGCGTGCACTGTGTCCG---ACCCCC----- 1203
 QY 669 TyrArgThrValTyrGlyGlnGlyAspGlySerIleGlyProIleGlySerThrArg 688
 Db 1204 -----GACTCTGGGCTAGTG-----TGTGGCCGC 1227
 QY 689 MetProAspAlaValThrHis-----HisThrCysIleLysSerThr 703
 Db 1228 ATGGCTGCTCCCGACGACGATGAGCAGCTCTGACGCGGGGCAACCCCGAGGGG 1287
 QY 704 GlutyrGluLeuGlyLeuLeuLysGluIleArgPheIleThrAspGlnLeuArgLys 723
 Db 1288 GACCCGAGCTTGGCCAGATCTCTGGAGAGTCCGCTACATTGCCAATCGCTTCGCTGC 1347
 QY 724 AspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArg 743
 Db 1348 CAGGACGAAAGCGAGGCGTCTGACGAGTGAAGTTCGCCGCTGTGTGGTGGACGC 1407
 QY 744 LeuCysIleIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuSer 763
 Db 1408 CTGTGCTCATGGCTTCTGGGTCTTCCATCATCTGCACCATCGGCATCTCTGATGTCG 1467
 QY 764 AlaProHisIleIle 768
 Db 1468 GCTCCCACTTCGTG 1482
 RESULT 13
 AAV12197
 ID AAV12197 standard; cDNA; 1876 BP.
 XX
 AC AAV12197;
 XX
 DT 14-MAY-1998 (first entry)
 XX
 DE Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.
 KW Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
 KW brain tissue; screening; NACHR; antibody; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 73..1581
 CDS /tag= a
 FT /product= "neuronal nicotinic acetylcholine receptor
 FT alpha-7 subunit"
 XX
 PN W09420617-A2.
 XX
 PD 15-SEP-1994.
 XX
 PF 08-MAR-1994; 94MO-US002447.
 XX
 PR 08-MAR-1993; 93US-00028031.
 XX
 PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
 XX
 PI Elliott KJ, Ellis SB, Harpold MM;
 XX
 DR WPI; 1994-303024/37.
 DR P-PSDB; AAW44153.
 XX
 PT Human neuronal nicotinic acetylcholine receptor subunits and DNA - also
 PT transformed cells useful for screening cpds. which modulate activity of
 PT the receptor.
 XX
 PS Claim 8; Page 78-79; 99pp; English.
 XX
 CC The present sequence encodes a human neuronal nicotinic acetylcholine
 CC receptor (NACHR) subunit. The cells expressing the alpha and/or beta
 CC NACHR subunits may be used in a method of screening compounds to identify
 CC any which modulate the activity of human neuronal NACHR. Subunit specific
 CC antibodies may be used to monitor the distribution and expression density

CC of various subunits in normal vs diseased brain tissues. Testing of
 CC single receptor subunits or specific receptor subunit combinations with a
 CC variety of potential agonists or antagonists provides information with
 CC respect to the function and activity of the individual subunits and
 CC should lead to the identification and design of compounds that are
 CC capable of very specific interaction with one or more receptor subtypes.
 CC The resulting drugs should exhibit fewer unwanted side effects than drugs
 CC identified e.g. screening with cells that express a variety of subtypes
 XX
 SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 1,14e-94 Length: 1876
 Score: 1159.50 Matches: 237
 Percent Similarity: 60.76% Conservative: 82
 Best Local Similarity: 45.14% Mismatches: 115
 Query Match: 28.68% Indels: 91
 DB: 2 Gaps: 12

US-09-303-232-2 (1-770) x AAV12197 (1-1876)

QY 297 IleTyrLeuAsnLeuSerAlaLys-----ValCysLeuAlaGlyTyrHisGluLys 313
 Db 94 GTCTGCTGGCGCTGGCCGGTGGCTCTCTGACGTGTCTCTGCAAGCGAGTTCAGAGG 153
 QY 314 ArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGlu 333
 Db 154 AAGCTTTACAAGGAGCTGTGCAAGAACTACAATCCCTTGAGAGGCGCTGGCCAATGAC 213
 QY 334 SerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGlu 353
 Db 214 TCGAACCACTACCGCTCTCTCTCTGAGCTCTCTGAGATCATGGAGTGGAGAG 273
 QY 354 LysAsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeu 373
 Db 274 AAGAACCAGTTTAAACCACCAACATTTGGCTGCAAAATGCTGGACAGATCATATTTA 333
 QY 374 ArgTrpAsnThrSerAspTyrGlyValLysAspLeuArgIleProHisArgIle 393
 Db 334 CAGTGAATGTGTGAGAATATCCAGGGGTGAAGACTGTTCTGTTCCAGATGGCCAGATT 393
 QY 394 TrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGln 413
 Db 394 TGGAAACCAAGATTTCTCTATAACAGTGTGATGAGCGCTTTGACGCCACATTCAC 453
 QY 414 ThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIlePheLys 433
 Db 454 ACTAAGTGTGTGTAATCTTCTGGGCATTTGCCAGTACCTGCTCCAGGCATATTCAAG 513
 QY 434 SerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLys 453
 Db 514 AGTTCTGTACATCATGATGCTGCTTCTCTCTGATGTGCGACACTGCAAACTGAAG 573
 QY 454 PheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGluThrGly 473
 Db 574 TTTGGTCTGTGTCTTACGAGGCTGTCTCTGGAATCTGCAGATGCAGAG----- 624
 QY 474 GlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLys 493
 Db 625 GCAGATATCAGTGGCTATATCCCAATGAGAATGGACCTAGTGGGAATCCCCGGCAAG 684
 QY 494 ArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIle 513
 Db 685 AGGAGTGAAGGTTCATGATGCTGCTCAAGAGAGCCCTACCCGATGTCACCTTCACAGTG 744
 QY 514 IleIleArgArgThrLeuTyrTyrPheAsnLeuIleIleProCysValLeuIle 533
 Db 745 ACCATCGCGGAGGAGCGCTCTACTATGGCTCAACCTGATCCCTGTGTGCTCATC 804
 QY 534 AlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeu 553
 Db 805 TCCGCCCTCGCCCTGTGTGTGTCTCTGCTCTCTGTCAGATTCGGGGAGAGATTTCCTCG 864

Db 94 GTCTGGCTGGCGTGGCGGCTGCTCTCTGTCACGTGCTCTGCAAGCGAGTTCACAGG 153
Qy 314 ArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGlu 333
Db 154 AAGCTTTACAGGAGCTGGTCAAGACTACAACTCCCTTGGAGAGCGCGTGGCCAAATGAC 213
Qy 334 SerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnLeuLeuLeuAspValAspGlu 353
Db 214 TCGCAACCACTCACCGTCTACTCTCTCCCTGAGCTCTCTGAGATCATGGAGTGGATGAG 273
Qy 354 LysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeu 373
Db 274 AAGAACAAGTTTAAACCAACCAATTTGGCTGCAATGTCTTGGACAGATCACTATTTA 333
Qy 374 ArgTrpAsnThrSerAspTyrGlyValLysAspLeuArgLeuProProHisArgGlu 393
Db 334 CAGTGGATGTGTCAGATATCCAGGGGTGAAGACTGTCTGTTTCCAGATGGCCAGATT 393
Qy 394 TrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGln 413
Db 394 TGGAAACCAAGACATCTCTCTATAACAGTGTCTGATGAGCGCTTTGACGCCACATCCAC 453
Qy 414 ThrAsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyLysPheLys 433
Db 454 ACTAACGTGTGGTGAATCTCTTGGGCATTGGCAGTACCTGCGCTCCAGGCATATCAAG 513
Qy 434 SerThrCysLysLeuAspLeuThrTrpPhePheAspAspGlnArgCysGluMetLys 453
Db 514 AGTTCCTGCTACATCGAATGACGTGTGTTTCCCTTTGATGTCAGCACTGCAAACTGAAG 573
Qy 454 PheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGly 473
Db 574 TTTGGTCTCTGCTTACGGAGGTGGTCTCTTGGATCTGCAGATGCAGGAG----- 624
Qy 474 GlyAspLysSerTrpValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLys 493
Db 625 GCAGATATCAGTGGCTATATCCCAATGGGAATGGGACCTAGTGGGAATCCCGGGAAG 684
Qy 494 ArgAsnGluLeuTyrTyrAsnCysCysProGluProTyrLeuAspLeuThrPheAlaLeu 513
Db 685 AGGATGAAAGTTCTATGAGTGTGCAAGAGCCCTACCCGATGTCACCTTCACAGTG 744
Qy 514 IleLeuArgArgThrLeuTyrTyrPhePheAsnLeuLeuLeuLeuProCysValLeuLeu 533
Db 745 ACCATGCGCGCAGACGCTCTACTATGGCTCAACCTGCTGATCCCTCTGTGCTCATC 804
Qy 534 AlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeu 553
Db 805 TCCGCCCTCGCCCTGCTGGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864
Qy 554 GlyValThrLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetPro 573
Db 865 GGGATAACAGTCTACTCTCTTACCGTCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
Qy 574 AlaThrSerAspAlaValProLeu----- 581
Db 925 GCAACATCCGATTCCGTTACCATTTGATAGCCAGTACTTGGCCAGCACCACCATCATCGTG 984
Qy 581 ----- 581
Db 985 GGCCTCTCGTGGTGGTGGATCGTGTGTCAGTACACACACACACACACACACACACAC 1044
Qy 582 ----- 582
Db 1045 GGCAAGATGCCAAGTGGACAGAGTCTCTTCTGAACTGGTGGCGGTGGTCTCTSCGA 1104
Qy 597 MetSerArgProGly-----ArgProLeuLeuLeuGluPheProThrThrPro 612
Db 1105 ATGAAGAGGCGCGGGGAGACAGAGTGGCGCCG---GCCTGCCAGACACAGCAGCGCGC 1161
Qy 613 CysSerAspThrSerGluArgLysHisGlnLeuLeuSerAspValGluLeuLysGlu 632
Db 1162 TGCAGCTTGGCAGTGTGGAG-----ATGAGCGCGCGTGGCGCGCGCGCGCGCG 1206

Qy 633 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspLeuAspAspPheArgHis--- 651
Db 1207 GCCAGAACGGGAACCTGCTGTAC-----ATCGGCTTCCGGCGGCTG 1248
Qy 652 -----AsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPhe 668
Db 1249 GACGGCGTGACATGTGTCCCG---ACCCCC----- 1275
Qy 669 TyrArgThrValTyrGlyGlnGlyAspAspGlySerLeuGlyProIleGlySerThrArg 688
Db 1276 -----GACTCTGGGTAGTG-----TGTGGCGCGC 1299
Qy 689 MetProAspAlaValThrHis-----HisThrCysIleLysSerSerThr 703
Db 1300 ATGCCCTGTCTCCCAACGACGATGACACCTCTCTGACGGCGGGCAACCCCGAGGGG 1359
Qy 704 GluTyrGluLeuGlyLeuLeuLeuLysGluLeuArgPheIleThrAspGlnLeuArgLys 723
Db 1360 GACCGGACTTGGCCAAAGATCTCTGGAGGAGTCCGCTACATTGCAATCGCTTCCGCTGC 1419
Qy 724 AspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArg 743
Db 1420 CAGGACGAAAGCGAGCGGTCTGCAGCGAGTGGAAAGTTCCGCCCTGTGTGGTGGACCGC 1479
Qy 744 LeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSer 763
Db 1480 CTGTGCTCTGAGGCTTCTCGGTCTTCACCATCATCTGCACCATCGGCATCTGATGTG 1539
Qy 764 AlaProHisIleIle 768
Db 1540 GCTCCCAACTTCGTG 1554

Search completed: May 8, 2004, 19:29:44
Job time : 850.783 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:31:03 ; Search time 18.738 Seconds

(without alignments)
3952.801 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKLTVEVDDDELAV.....MFAILLATIAVLLSAPHIIVS 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182.5	29.2	502	2 A57175	nicotinic acetylch
2	1179.5	29.2	502	2 JN0113	nicotinic acetylch
3	1173.5	29.0	502	2 T01378	nicotinic receptor
4	1168	28.9	511	2 JH0173	alpha-bungarotoxin
5	1160.5	28.7	502	1 ACHUA7	nicotinic acetylch
6	1155.5	28.6	502	2 G02259	alpha 7 neuronal n
7	1104	27.3	498	2 S68588	nicotinic acetylch
8	1009.5	25.0	461	2 T25671	hypothetical prote
9	988	24.4	560	2 T19622	hypothetical prote
10	968.5	24.0	542	2 T19862	hypothetical prote
11	889	22.0	567	1 ACEFA1	nicotinic acetylch
12	873	21.6	557	2 T12359	nicotinic acetylch
13	845	20.9	503	2 A53956	nicotinic acetylch
14	843.5	20.9	576	1 ACPEA2	nicotinic acetylch
15	841	20.8	521	1 ACEFNN	nicotinic acetylch
16	840	20.8	502	2 A37040	nicotinic acetylch
17	838	20.7	528	1 ACCH2N	nicotinic acetylch
18	828	20.5	511	2 A40110	nicotinic acetylch
19	827.5	20.5	456	1 ACCHAN	nicotinic acetylch
20	824	20.4	499	2 A24572	nicotinic acetylch
21	820.5	20.3	495	2 S60589	acetylcholine rece
22	820.5	20.3	512	2 B37014	nicotinic acetylch
23	817	20.2	494	2 T09289	nicotinic acetylch
24	806.5	19.9	500	2 T12899	nicotinic acetylch
25	805	19.9	498	2 G02421	nicotinic acetylch
26	802.5	19.8	457	1 ACHUA1	nicotinic acetylch
27	801	19.8	457	1 ACBOA1	nicotinic acetylch
28	794.5	19.7	457	2 A28529	nicotinic acetylch
29	794	19.6	457	2 A24383	nicotinic acetylch

ALIGNMENTS

RESULT 1

A57175

nicotinic acetylcholine receptor alpha-7 chain precursor, neuronal - mouse
C:Species: Mus musculus (house mouse)

C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 20-Aug-1999

C:Accession: A57175

R:Orr-Urtreger, A.; Seldin, M.F.; Baldini, A.; Beaudet, A.L.

Genomics 26, 399-402, 1995

A:Title: Cloning and mapping of the mouse alpha7-neuronal nicotinic acetylcholine recep

A:Reference number: A57175; MUID:95324936; PMID:7601470

A:Accession: A57175

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-502 <ORR>

A:Cross-references: GB:L37663; NID:q790853; PIDN:AAC42053.1; PID:q790854

C:Superfamily: acetylcholine receptor

C:Keywords: brain; glycoprotein; ion channel; neurotransmitter receptor; phosphoprotein

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status pred

F:231-254/Domain: transmembrane #status predicted <TR1>

F:262-280/Domain: transmembrane #status predicted <TR2>

F:296-317/Domain: transmembrane #status predicted <TR3>

F:470-488/Domain: transmembrane #status predicted <TR4>

F:46,90,133/Binding site: carboxylate (Asn) (covalent) #status predicted

F:365,413,427/Binding site: phosphate (Ser) (covalent) #status predicted

F:415/Binding site: phosphate (Thr) (covalent) #status predicted

F:442/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 29.2%; Score 1182.5; DB 2; Length 502;

Best Local Similarity 45.8%; Pred. No. 5e-77;

Matches 239; Conservative 80; Mismatches 118; Indels 85; Gaps 10;

QY 297 IYLNLSA---KVCILAGYHEKRLHLLDPYNTLERPVNLSPEQLQSLGLTQIWDVDE 353

DB 8 IWLAAALLHVSLOGEFQRLYKELVKNYNPLRPVANDSQPLTVYFSLSLQIMDVE 67

QY 354 KNQLIVTNVLKLEWNNMNLNRTSDYGGVKDLRIPPHRIWKFPDVLNNSADEGDTGYQ 413

DB 68 KNOVLTNNILQMSWTDHYLQNNMSEYGVKNVRFPDQGIWKFPDILLNNSADERFDTFH 127

QY 414 TNVVRNNGSLCYVDPGIFKSTCKIDITWFFPDQRCMKGSGWTYDGFQDLQLODRTG 473

DB 128 TNLVNASGHQYLPFGIFKSCYIDVRVFPDVQOQCKLFGSGWSYGGSLDLQMQE--- 184

QY 474 GDISSYVLNGWELLGVGKNEIYNCCPEYDITFALLIRRTLYFFNLLIPCVLI 533

DB 185 ADISSYIPNGWDLMGIPGKNEXPYDVTVTMRRRTLYYGLNLLIPCVLI 244

QY 534 ASMALLGFTLPDPSGEKLSLGVITLLSLTVFLNVAETMPATSDAVPL----- 581

DB 245 SALALLVFLLPADSGEKLSLGLTVLLSLTVFMLLVAEIMPATSDVPLIAQYFASTMLIV 304

acetylcholine rece
nicotinic acetylch
nicotinic acetylch
nicotinic acetylch
nicotinic acetylch
acetylcholine rece
nicotinic acetylch
nicotinic acetylch
protein F25G6.4 [i
nicotinic acetylch
nicotinic acetylch
hypothetical prote
nicotinic acetylch
nicotinic acetylch
hypothetical prote

QY 582 -----WIRIVFLCWLPMILMSRPG-----RPLILEPPTTP 612
Db 305 GLSVVTVVILVYHHDDPGGKMPKWTIRILLNWCWFLMKPGEKVKVRPACQHKPRR- 363
QY 613 CSDTSSERKHOILSDVELKERSKSLANVLDDIDDFRHNCRPMTPGGLPHNPAPYRTV 672
Db 364 CS-----LASVELSAGAG-----PPTSGNLLY--IGFRGL 392
QY 673 YQG-----GDDSGI-SPIGSTRMPDAVTHHTCIKSTVEYELGLILKEIRFITDQLRKDDE 726
Db 393 EGMHCAPTDSGVVVCGLACSPTHDEHLMHGTHSPDGLAKILBEVRYIANRFRQCODE 452
QY 727 CNDIANDWKAAMVVDRLCLIIFTMEAILATIAVLLSAPHII 768
Db 453 SEVICSEWFAACVVDRLCLMAFSVFTICTIGILMSAPNFV 494
RESULT 2
JN0113
N;Alternate names: alpha-bungarotoxin-binding protein alpha chain - chicken
C;Species: Gallus gallus (chicken)
C;Date: 12-Feb-1993 #sequence revision 12-Feb-1993 #text_change 20-Aug-1999
C;Accession: JN0113; JH0172; S28018; B25738; S26566
R;Couturier, S.; Bertrand, D.; Matter, J.M.; Hernandez, M.C.; Bertrand, S.; Millar, N.;
Neuron 5, 847-856, 1990
A;Title: A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is developmentally
A;Reference number: JN0113; MUID:91097796; PMID:1702646
A;Accession: JN0113
A;Molecule type: DNA
A;Residues: 1-502 <COU>
A;Cross-references: GB:X68586; NID:9287756; PIDN:CAA48576.1; PID:g287757
A;Experimental source: white leghorn; brain
R;Schoepfer, R.; Conroy, W.G.; Whiting, P.; Gore, M.; Lindstrom, J.
Neuron 5, 35-48, 1990
A;Title: Brain alpha-bungarotoxin binding protein cDNAs and Mabs reveal subtypes of this
A;Reference number: JH0172; MUID:90315158; PMID:2369519
A;Accession: JH0172
A;Molecule type: mRNA
A;Residues: 1-502 <SCH>
A;Cross-references: EMBL:X52295; NID:g63077; PIDN:CAA36543.1; PID:g63078
A;Experimental source: brain
R;Matter-Sadzinski, L.; Hernandez, M.C.; Roztocil, T.; Ballivet, M.; Matter, J.M.
EMBO J. 11, 4529-4538, 1992
A;Title: Neuronal specificity of the alpha7 nicotinic acetylcholine receptor promoter de
A;Reference number: S28018; MUID:93049204; PMID:1425587
A;Accession: S28018
A;Molecule type: DNA
A;Residues: 1-18 <MAS>
A;Cross-references: EMBL:X68246; GB:949751; NID:g65319; PIDN:CAA48317.1; PID:g65320
A;Experimental source: white leghorn; erythrocyte
R;Conti-Tronconi, B.M.; Dunn, S.M.J.; Barnard, E.A.; Dolly, J.O.; Lai, F.A.; Ray, N.; Ra
Proc. Natl. Acad. Sci. U.S.A. 82, 5208-5212, 1985
A;Title: Brain and muscle nicotinic acetylcholine receptors are different but homologous
A;Reference number: A94055; MUID:85270494; PMID:3860855
A;Accession: B25738
A;Molecule type: protein
A;Residues: 24-25, 'EN', 28-41, 'X', 43-45, 'X', 47 <CON>
C;Comment: This acetylcholine receptor is blocked by alpha-bungarotoxin and is localized
C;Genetics:
A;Introns: 19/1; 65/3; 80/3; 117/2; 144/1; 200/1; 265/1; 294/1; 330/3
C;Superfamily: acetylcholine receptor
C;Keywords: brain; glycoprotein; ion channel; neurotransmitter receptor; phosphoprotein;
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status predi
F;231-254/Domain: transmembrane #status predicted <TR1>
F;262-280/Domain: transmembrane #status predicted <TR2>
F;296-317/Domain: transmembrane #status predicted <TR3>
F;470-488/Domain: transmembrane #status predicted <TR4>
F;46, 90, 133/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;365, 367, 413, 427, 465/Binding site: phosphate (Ser) (covalent) #status predicted
F;415/Binding site: phosphate (Thr) (covalent) #status predicted
F;442/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 29.2%; Score 1179.5; DB 2; Length 502;
Best Local Similarity 45.6%; Pred. No. 8.2e-77;
Matches 237; Conservative 80; Mismatches 127; Indels 75; Gaps 9;
QY 295 LIYINLSA---KVLGAGYHEKRLHDLDPYNTLIERPVLNBSDPQLQSLFGLTLMQIDV 351
Db 6 LMLLILAAAGLVRESLQGEFQKLYKELLKNTNPLRPVANDSQPLTVVFTLSLQIMDV 65
QY 352 DEKNQLVTNVWLKLEWDMNLRWNTSDYGGVKDLRIIPHRITWKPDVLMYNSADEGPDGT 411
Db 66 DEKNQVLTNIWLQYWTDHYLQWNVSEYPGVKNVRFDPGLIWKPDILLYNSADERDAT 125
QY 412 YOTNVVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGFQDLQLQDE 471
Db 126 FHTNVLNNSGHCQYLPFGIFKSSCYIDVRWFFDVQKCNKFGSWTYGWSLDLQME- 184
QY 472 TGGDISYVINGEWELLVPGKRNELIYNCCBPYDITFAIIRRTLYYFNLIIPCV 531
Db 185 --ADISGYISNGEWDLVGIPGKRTESFYECCKEYPDITFTVTRRTLYYGLNLLIPCV 242
QY 532 LIASMLLGLTLPDPSGEKLSLGTLLISLTVFLNVAETMPATSDAVPL- 581
Db 243 LISALALLVFLPADSGEKLSLGTIVLLSLTVFMLLVAEIMPAISVPLAQYFASTMI 302
QY 582 -----WIRIVFLCWLPMILMSRPGRLPILFPTTPCS 614
Db 303 IVGLSVTVTVIVLYHHDDPGGKMPKWTIRILLNWCWFLMKRPG-----EDKVRPAC 357
QY 615 DTSSERKHOILSDVELKERSKSL-ANVLDDIDDFRH-----NCRPMTPGGLPHNPAPY 669
Db 358 QHKQRCS--LSSMEMNTVSGQCSNGNMLYI--GFRGLDGVRCTPTDTSVGI- 406
QY 670 RTVYGGDDSGIGPIGSTRMPDAVTHHTCIKSTVEYELGLILKEIRFITDQLRKDDECD 729
Db 407 -----CGRMTCSPTEENLHSGHPGSDPLAKILBEVRYIANRFRDQDEEA 455
QY 730 IANDWKAAMVVDRLCLIIFTMEAILATIAVLLSAPHII 768
Db 456 ICNEWKAAMVVDRLCLMAFSVFTICTIGILMSAPNFV 494
RESULT 3
T01378
nicotinic receptor alpha 7 chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jul-2000
C;Accession: T01378
R;Seguela, P.; Wadiche, J.; Dineley-Miller, K.; Dani, J.A.; Patrick, J.W.
J. Neurosci. 13, 596-604, 1993
A;Title: Molecular cloning, functional properties, and distribution of rat brain alpha
A;Reference number: Z14310; MUID:93147931; PMID:7678857
A;Accession: T01378
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-502 <SEG>
A;Cross-references: EMBL:S53987; NID:g264770; PIDN:AAB25224.2; PID:g5705903
A;Experimental source: brain
C;Superfamily: acetylcholine receptor

Query Match 29.0%; Score 1173.5; DB 2; Length 502;
Best Local Similarity 45.6%; Pred. No. 2.2e-76;
Matches 238; Conservative 80; Mismatches 119; Indels 85; Gaps 10;
QY 297 IYINLSA---KVLGAGYHEKRLHDLDPYNTLIERPVLNBSDPQLQSLFGLTLMQIDVDE 353
Db 8 IWLALAAALHVLQGEFQRLYKELVKNVPLRPVANDSQPLTVVFTLSLQIMDVDE 67
QY 354 KNQQLVTNVWLKLEWDMNLRWNTSDYGGVKDLRIIPHRITWKPDVLMYNSADEGFDGTQY 413
Db 68 KNQVLTNIWLQSWTDHYLQWNVSEYPGVKNVRFDPGQIWKPDILLYNSADERFDATFH 127
QY 414 TNNVVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGFQDLQLQDETG 473

```

Db 128 TNLVNASGHCQVLPGLKSSCYIDVRWFDPVQCKLFGSWSYGGSLDLQME---184
QY 474 GDSSVVLNGEWELLGVPKRNELIYNCCPEYIDITFAIIRRTLYFFNLIIPCVLI 533
Db 185 ADISSYIPNGEWMIDMGPCKRNEKFEYCKPEYDPVTVYTWRRRTLYYGLNLLIPCULI 244
QY 534 ASWALLGFTLPDPSGKLSIGVTLLSLVFLNVAETMPATSDAVPL-----581
Db 245 SALALLVFLPADSGEKISLGITVLLSLVFLMLVAEIMPATSDSVPLIAQVASTMIIV 304
QY 582 -----WIRIVFLCWLFWILMSRPG-----RPLILEPPTP 612
Db 305 GLSVVVTVIVLRVHHHDPDGKMPKWTRIILLNKCWAFLEMKPEGDVKRPACQKPRR- 363
QY 613 CSDTSERKHQILSDVELKERSKSLIANVLDDDDFRHNCRPMTPGGTLPHNPAFYRV 672
Db 364 CS-----LASVELSAGAG-----PPTSGNLLY--IGFRGL 392
QY 673 YGQ-----GDGSI-GPIGSTRMPDAVTHHTCTKSSTEYELGLILKEIRFITDOLKXDE 726
Db 393 EGHWCAPTDSGVVCGRLACSPTHDHLMHGAHPSDGDPLAKILLEEVRYIANRNCODE 452
QY 727 CNDIANDKFAAMVVDRLCLIIFTFPAIATIAVLLSAPHII 768
Db 453 SEVICSEWKAACVVDPLCLMAFSVFTICTIGILMSAPNFV 494

RESULT 4
JH0173
alpha-bungarotoxin-binding protein alpha-2 chain precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 20-Aug-1999
C:Accession: JH0173
R:Schoepfer, R.; Conroy, W.G.; Whiting, P.; Gore, M.; Lindstrom, J.
Neuron 5, 35-48, 1990
A:Title: Brain alpha-bungarotoxin binding protein cDNAs and MABs reveal subtypes of this
A:Reference number: JH0172; MUID:90315158; PMID:2369519
A:Accession: JH0173
A:Molecule type: mRNA
A:Residues: 1-511 <SCH>
A:Cross-references: GB:X52296; NID:963081; PIDN:CAA36544.1; PID:G63082
A:Experimental source: brain
A:Note: this sequence is similar to acetylcholine receptor alpha chains
C:Comment: Alpha-bungarotoxin binding proteins are localized to extrasynaptic pseudodend
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-510/Product: alpha-bungarotoxin binding protein alpha-2 chain #status predicted <AR
F:239-262/Domain: transmembrane #status predicted <TM1>
F:270-288/Domain: transmembrane #status predicted <TM2>
F:304-323/Domain: transmembrane #status predicted <TM3>
F:479-496/Domain: transmembrane #status predicted <TM4>
F:54/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.9%; Score 1168; DB 2; Length 511;
Best Local Similarity 44.2%; Pred. No. 5.6e-76;
Matches 238; Conservative 83; Mismatches 117; Indels 100; Gaps 12;

QY 285 NGLNKHSTWFLIYILNLSAKVCLAGVHEKRLSLHLLDPYNTLPERPVNLSDEPLQLSFGIT 344
Db 12 SGLCLWASLFTGFF-----KVSGQGESQRRLYRDLRLNRYNLERPWNDSQIVVELQLS 66
QY 345 LMQIIDVDEKQLLVTVNWVLEWDMNLRWNTSDYGVKDIIRIPHRIRKPDVLMYNSA 404
Db 67 LQQIIDVDEKQVLITNAWLOWYWDIYLSVDQYEPGVQNLRFPSDQIWIVPDILLXNSA 126
QY 405 DSGFDGTQTVNVRNNGSLVVPBGIFKSTCKIDITWPPDDQRCSEMKFGSWYDGLQL 464
Db 127 DERFATHTNVLVNVGSCQVIPPGLKSTCYIDVRWFPDFVQCKDLFGSWTHSGWLI 186
QY 465 DIQLQDETGGDISSVLNGEWELLGVPKRNELIYNCCPEYIDITFAIIRRTLYYFF 524

```

```

Db 187 DLQMLE---ADISNYISNGEWDLVGVPGKRNELIYECCKPEYDPVTVYTTMRRRTLYYGL 243
QY 525 NLIIPCVLIASMALLGFTLPDPSGKLSIGVTLLSLVFLNVAETMPATSDAVPL---581
Db 244 NLLIPCVLISGLALLVFLPADSGEKISLGITVLLSLVFLMLVAEIMPATSDSVPLIAQ 303
QY 582 -----WIRIVFLCWLFWILMSRPG-----RPL 604
Db 304 YFASIMVIVGLSVVVTVLVLQFHHHDPQAGKMPWRVVRVILLNKCWAFLEMKKPGENIKPL 363
QY 605 ILEFPTTSCDTSERKHQILSDVELKERSKSLIANVLDDDDFRHNCRPMTPGGTLPH 664
Db 364 SCKY-----SYPKHPSLKNTEM-----NVL-----PG---H 387
QY 665 NPAFVTVVYG-----QGDD--GSIGPIGSTRMPDAVTHHTCTKS--STEYELGLI 710
Db 388 QPSNGNMIYSYHTMENPCPPQNLDLGSKGKITPLSED--NEHVQKXALMDTIPVVKI 445
QY 711 LKEIRFITDOLKDECDNDIANDKFAAMVVDRLCLIIFTFPAIATIAVLLSAPHII 768
Db 446 LEEVOFIAMRFRKQDEGEIEICSEWKAACVVDRLCLVAFITLFAICTFTILMSAPNFI 503

RESULT 5
ACHUA7
nicotinic acetylcholine receptor alpha-7 chain precursor, neuronal - human
N:Alternate names: cholinergic nicotinate receptor alpha-7 chain
C:Species: Homo sapiens (man)
C>Date: 12-Aug-1996 #sequence_revision 31-Jan-1997 #text_change 22-Jun-1999
C:Accession: I37185; A54194; S60309
R:Peng, X.; Katz, M.; Gerzanich, V.; Anand, R.; Lindstrom, J.
Mol. Pharmacol. 45, 546-554, 1994
A:Title: Human alpha 7 acetylcholine receptor: cloning of the alpha 7 subunit from the
omers expressed in Xenopus oocytes.
A:Reference number: I37185; MUID:94195283; PMID:8145738
A:Accession: I37185
A>Status: translated from GB/EMBL/DBSJ
A:Molecule type: mRNA
A:Residues: 1-502 <PEN>
A:Cross-references: EMBL:X70297; NID:9496606; PIDN:CAA49778.1; PID:g496607
A:Experimental source: brain neuroblastoma cell line SHSY-5Y
R:Chini, B.; Raimond, E.; Elgoyhen, A.B.; Moralli, D.; Balzaret, M.; Heinemann, S.
Genomics 19, 379-381, 1994
A:Title: Molecular cloning and chromosomal localization of the human alpha-7-nicotinic :
A:Reference number: A54194; MUID:94245214; PMID:8186270
A:Accession: A54194
A:Molecule type: mRNA
A:Residues: 24-363, 'S', 365-374, 'A', 376-408, 'AWPAP', 414-502 <CHI>
A:Cross-references: GB:Z23141; NID:g457736; PIDN:CAA80672.1; PID:g457737
A:Experimental source: retina
C:Comment: This acetylcholine receptor is blocked by alpha-bungarotoxin and is localized
C:Genetics:
A:Gene: GDB:CHRNA7
A:Cross-references: GDB:I38751; OMIM:118511
A:Map position: 15q14-15q14
A:Note: defects in this gene have been associated with mental retardation and schizophre
C:Complex: the functional receptor molecule is a heteropentamer with two alpha chains ar
C:Superfamily: acetylcholine receptor
C:Keywords: brain; glycoprotein; heteropentamer; ion channel; neurotransmitter receptor;
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status predi
F:231-254/Domain: transmembrane #status predicted <TR1>
F:262-280/Domain: transmembrane #status predicted <TR2>
F:296-317/Domain: transmembrane #status predicted <TR3>
F:470-488/Domain: transmembrane #status predicted <TR4>
F:46.90.133/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:150-164/Disulfide bonds: #status predicted
F:365.413/Binding site: phosphate (Ser) (covalent) #status predicted
F:365.413/Binding site: phosphate (Thr) (covalent) #status predicted
F:442/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 28.7%; Score 1160.5; DB 1; Length 502;
Best Local Similarity 45.1%; Pred. No. 1.9e-75;

```

```

185 ADISGVIPNGEWDLVIGPGKRSEFYECCKEYPEDVTFVTVMRRRTLYGLNLLIPCVLII 244
QY 534 ASMALLGFTLPDPSGEKLSIGVILLSLTVFLNNVAETMPATSDAVPL----- 581
Db 245 SALALLVFLLPADSGEKISLGITVLLSLTVFMLLVAEIMPATSDSVPLIIAQVFASFTMIIV 304
QY 582 -----WTRIVFLCWLPMILRMSRPG-----RPLILEFPPTP 612
Db 305 GLSVWVTIVLQYHHDDPDGCKMPKTRVILLNWCWFLMRKRPGECDKVRP-ACQHKORR 363
QY 613 CSTSSSERKHQIILSDVELKERSKSLIANVLIDDDDERH-----NCRMTPGGTLPHNPAF 668
Db 364 CSLASVE-----NSAVAPPPASNGLLY-----IGRGLDGVHCVP-Tp----- 401
QY 669 YRTVYGGDDSGTGPIGSTGMPDAVTH-----HFCIKSSTYEYELGLILKEIRFITDQLRK 723
Db 402 -----DSGVV-----CGRMACSPTHDEHLLHGGQPPGDPDLAKILEEVRYIANRPRC 449
QY 724 DDECNPIANDKWFAMVVDRLCLIIFTWFAILATIAVLLSAPHII 768
Db 450 QDESEAVCSWKFAACVVDRLCLMFASFVTIICTIGILMSAPNFV 494

RESULT 7
S68588
nicotinic acetylcholine receptor alpha-1 chain precursor (clone Ce21) - Caenor
C;Species: Caenorhabditis elegans
C;Date: 06-Dec-1996 #sequence_revision 07-Feb-1997 #text_change 20-Aug-1999
C;Accession: S68588; S57496
R;Ballivet, M.; Alliod, C.; Bertrand, S.; Bertrand, D.
J. Mol. Biol. 258, 261-269, 1996
A;Title: Nicotinic acetylcholine receptors in the nematode Caenorhabditis elegans
A;Reference number: S68587; MUID:96196478; PMID:8627624
A;Accession: S68588
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-498 <BAL>

```

C; Keyword

```
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-498/Product: nicotinic acetylcholine receptor alpha-1 chain #status predi
```

	Query Match	27.3%	Score 1104;	DB 2;	Length 498;	
	Best Local Similarity	43.3%;	Pred. No. 2.1e-71;			
	Matches 227;	Conservative	76;	Mismatches 139;	Indels 82;	Gaps
QY	295	LLIVNLISAKVCLAGYHEKHLHDLDPYNTLERPVLNESPQLQSFGLTLMQIIDVDEK	354			
Db	6	LLISCAILAAPTIGSQERRLIEDLNRNNYLNPVANHSPEVTVHLKVAQQIIDVDEK	65			
QY	355	NQLLVTVNVKLKWMDMNLRWNTSDYGGVKDRLRIPPHRIWKPDVLMYNASDBGFTGYQT	414			
Db	66	NQVVYNAWLDYTWNLYNDVKAEYGNITDVRFPAGIKMPDVLLXNSVDTNFSTGYT	125			
QY	415	NVVVRNGSGLYPPPGIFKSTCKIDIITWPPDDQCCEMKFGSWTDGFDLDLOODETCG	474			
Db	126	NMIYSTGLVHWPPPGIFKISKIDIQWFPEDEQCFKFGSWTDGFKLD--LPATGG	183			
QY	475	-DISSVYLVNGEWELLGVPGKRNEYNYNCPEYIDITPAIIIRRRRTLYFFNLIIIPCVI	533			
Db	184	FDISEYSNGEAWPLTTVERNEKEFYDCPEPYDVHFVHMRRTLYYGNLIMPCILT	243			
QY	534	ASWALLGFTYPPDSGEKLSGVTLILLSLVFLNMVAETMPATSDAVPI-----	581			
Db	244	TMLTLGFTLPPDAGEKITLIQTIVLLSICFELSIIVSEMSPTSEAVPLLGIHFFTCMIV	303			
QY	582	-----WRIVFCLMWILTRMSRGPURPILBPTTPCSDT	616			
Db	304	TASTVTFVYVNLNHYRTPETHDMCPWTENLLYYIWILLRMKRPGHN--LIYASLPSFS	361			
QY	617	SSEKKHQILDVELKERSSKLLANVID-----IDDFRNCRPMTPGGTLPHNP	666			

```
Db 362 TKPNRH-----SESLRNKNEHLSRANSFADCRNLQYIMTQSVS----- 404
QY 667 AFYRTYGGDDG--SIGPIGSTRMPDAVTHHTCTKSSTYELGLILKEIRFITDQLRKD 724
Db 405 -----NGLTSLGSPSTMISSNGTTDVSQATLLILHRIYHELKIVTKRMIEG 453
QY 725 DRCNDIANDWKFAAMVVDRLCLIIPTMFALATIATVLLSAPHII 768
Db 454 DKEBOACNNKFAAMVVDRLCLYVFTIFIVSTIGIFWSPAPVIV 497

RESULT 8
T25671
hypothetical protein D2092.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T25671
R:Gattung, S.; Maggii, L.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid D2092.
A:Reference number: Z20067
A:Accession: T25671
A>Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-461 <GAT>
A:Cross-references: EMBL:U88167; PIDN:RAB42223.1; GSPDB:GN00019; CESP:D2092.3
A:Experimental source: strain Bristol N2; clone D2092
C:Genetics:
A:Gene: CESP:D2092.3
A:Map position: 1
A:Introns: 36/1; 80/3; 119/2; 191/1; 243/1; 274/1; 303/1; 382/2
C:Superfamily: acetylcholine receptor

Query Match 25.0%; Score 1009.5; DB 2; Length 461;
Best Local Similarity 39.4%; Pred. No. 1.1e-64;
Matches 209; Conservative 80; Mismatches 126; Indels 115; Gaps 10;

QY 283 ILNGLKHSWIFLLIYLNLSAKVCLAGYHEKRLHLLDLPYNTLPRVLNESDPIQLSPG 342
Db 2 IFNLNS-----ILSLVIHSLNCGSVAETKLTLDLKGYNLPRPVNSOPLEVTK 56
QY 343 LPLMQLIDVDEKNQLLTVTNVWKLWNMDNLRWNTSDYGGVXDLRIP--PHRIKPDVLM 400
Db 57 LFLQQLIDVDEKNQIVSNWLSYTFWHLQWEPKYGIGQIDIFPGSSDHIWKPDLV 116
QY 401 YNSADEGEGDTQTNVVRNNGSLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYD 460
Db 117 YNSAAEEDFSTPKSNLLTYHTGTVMVIPPGLKFCVQLDVTWPFDDQRCMKFGSWTFH 176
QY 461 GFOLDLQLODETGG---DISSVVLNGEWELLGVCPKNEIYVNCPEYDITFAIIR 516
Db 177 GYALDQLDDDDNGTQSDMLSTYLWNGEQQVISTNAKRVSYKCCPEPTVNYLH 236
QY 517 RRTLYFFNLIIIPCVLIASMAILGFTLPDPSGKLSGLVTILLSTVFLNVAETMPATS 576
Db 237 RRTLYYGNLIIIPSLIISMAILGFMFPDPAKEKITLEVITLLAIVFLSWVSEMTPTS 296
QY 577 DAVPL-----WIRIVFLCWLPWILMSR 599
Db 297 EAVPLIGVFFSCMLWSASVVFTIIVLNLHFRSADSHEMNLVRRVLLLEFLPWLNSR 356
QY 600 PCRPLILEPTTSCDTSSTSRKHQILSDVELKERSKSLANVLIDDDDFRNCRMP- 658
Db 357 PG-----YKFVKANVIDSTDKMPK--RPKNPL 381
QY 659 GGTLPHPNPAFYRTYVYGGDGIPIGSTRMPDAVTHHTCTKSSTYELGLILKEIRFIT 718
Db 382 DCNLPNSHAGY-----EAQILLHLSV-----HT-----ELRRV 410
QY 719 DQLRKDDCNDIANDWKFAAMVVDRLCLIIPTMFALATIATVLLSAPHII 768
Db 411 AFYNKEEHDRITQDWRFAAMVVDRLCLLTFTVFIIVISILAIMMSAPHII 460
```

RESULT 9

T19622
hypothetical protein C31H5.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T19622
R:Kershaw, J.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z19153

A:Accession: T19622

A>Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 1-560 <WIL>

A:Cross-references: EMBL:Z93778; PIDN:CAR07843.1; GSPDB:GN00019; CESP:C31H5.3

A:Experimental source: clone C31H5

C:Genetics:

A:Gene: CESP:C31H5.3

A:Map position: 1

A:Introns: 24/1; 70/3; 139/2; 270/1; 299/1; 336/3; 372/2; 456/3

C:Superfamily: acetylcholine receptor

Query Match 24.4%; Score 988; DB 2; Length 560;

Best Local Similarity 37.2%; Pred. No. 5.2e-63;

Matches 214; Conservative 85; Mismatches 152; Indels 124; Gaps 13;

QY 293 IFLLIYLNLSAKVCLAGYHEKRLHLLDLPYNTLPRVLNESDPIQLSFLTLMQIIDV 352

Db 12 VSILIWETKCSKVIWTGDHERLYAKLAENYNKLARPVNESEAVVLLGMDYQQIILDID 71

QY 353 EKNQLLVNVLKLEWMDNLRWNTSDYGGVXDLRIPPHRIWKPVDVLMVNSADEGPDGY 412

Db 72 EXHQIMNSNVLMRWSMTDHYLTWDPSEFGNIKEVRLPINNIWKPVDVLLNSVDQPDSTM 131

QY 413 QTNVVRNNGSLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDQFQLDQLQDET 472

Db 132 PVNAVLYTGNVTWIPPAIRSSCAIDTAYFFDQHCTMKFGSWTYSGFFIDL---INT 188

QY 473 GGDISSVVLNGEWELLGVCPKNEIYVNCPEYDITFAIIRRTLYFFNLIIIPCVL 532

Db 189 TISPATYKPNGEWELLGLTSQRSIFFECCPEPYDVTFTVSIIRRTLYYGNLIIIPCVL 248

QY 533 TASMAILGFTLPDPSGKLSGLVTILLSTVFLNVAETMPATSADVLWIRIVFLC--- 589

Db 249 ISSLALLSFTLPADCGEXNLGVTFMSLCVPMIWAEMAMPQTSALPL-IQIFYSCINF 307

QY 590 -----WLPWILMSRPGRPLILEPTTTC 613

Db 308 QVGASVATVIALNFHRSBPQYKPMNKLTKLLGLWLTLLGMRPD---VLELSVHGA 364

QY 614 SDTSERKHQ-----IL-----SDVELKERSKSL 638

Db 365 HYASDNKKKQRYLIEVERHILTRPNGNHSADVDAVHLDLSTGNPHSDAKSSSPKRT 424

QY 639 LANVLDD-----IDDDF-RHNC-----RPMTPG-GTLPH-----NPAPFRTYVG 674

Db 425 SASIMGMTGLPTQMGALDSSINKYCTKTVRPLENGSATINHKSSPOINNNNIYK 484

QY 675 QDDGSGIGPIGSTRMPDAVTHHTCTKSSTYELGLILKEIRFITDQLRKDDCNDIANDW 734

Db 485 CANN-----KQTFQEDRHFFH-----INELAVISARVKEEAMHALQADW 525

QY 735 KFAAMVVDRLCLIIPTMFALATIATVLLSAPHII 769

Db 526 MFASRVVDRCVCLAPSAFLMCTAISYNAPHLFV 560

RESULT 10

T19862
hypothetical protein C40C9.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T19862

A:Cross-references: EMBL:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576
A:Note: 538-Tyr was also found
C:Genetics:
A:Gene: FlyBase:nAcr-alpha-96Aa
A:Cross-references: FlyBase:FBN0000036
A:Map position: 3R 96A
A:Introns: 64/3; 79/3; 116/2; 176/3; 330/2; 401/1; 499/3
C:Superfamily: acetylcholine receptor
C:Keywords: Glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane;
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-567/Product: nicotinic acetylcholine receptor alpha-like chain #status predicted <SM>
F:22-240/Domain: extracellular #status predicted <EXT>
E:240-264/Domain: transmembrane #status predicted <TM1>
E:272-290/Domain: transmembrane #status predicted <TM2>
F:306-325/Domain: transmembrane #status predicted <TM3>
F:326-513/Domain: intracellular #status predicted <INT>
F:514-532/Domain: transmembrane #status predicted <TM4>
F:45,223/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:149-163,222-223/Disulfide bonds: #status predicted

```

F:306-323/Domain: transmembrane #status predicted <IN>
F:326-513/Domain: intracellular #status predicted <IN>
F:514-532/Domain: transmembrane #status predicted <TM>
F:45_223/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:149-163,222-223/Disulfide bonds: #status predicted

Query Match      22.0%  Score 889;  DB 1;  Length 567;
Best Local Similarity 35.2%;  Pred. No. 6.6e-56;
Matches 196;  Conservative 91;  Mismatches 168;  Indels 102;  Gaps 10;

Qy      291  SWIFLIYILNLS-AKVCLAGYHEKRLIHLDLPYNTLERPVLNESDPQLQSLGLTLMQII 349
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      3   SVLFAAVFIALHFATGGLANPDAKRLVYDILLSSNYNRLIRPVGNNSDRLTVMKGLRLSQLI 62

Qy      350  DVEKQQLLVNTWVWKLKWDMDNMLRWNTSDYGVKDLRIPEPHRIWPKDVLVYNSADEGFD 409
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      63  DVNLKQKQIMTNTVWVQEWMDYKLKNPDDYGGVDTLHVESEHIWHPDIYLVYNNADGNYE 122
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

Qy      410  GTYQTNTVVNRNNGSCILVPPSGIFKSTCKIDIITWFPDDQRCCKMFKFSWTYDGFQLDLI 466
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

696   HTCIKST-----EYELGLILKEIRFITDOLRKDDCECNDIANDWKFAAMVVDRLCLIIFT 750
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    461  AAADLSPTFEKPYAREMEKTIIEGSRFTAQHVKNKDKFESVEEDKKYAVMLDRMFLWIFA 520
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy    751  MFAILATIIVLLSAPHI 767
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    521  IACVVGTALIIQLAPSL 537
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
S12359
nicotinic acetylcholine receptor alpha-Li chain precursor - desert locust
C;Species: Schistocerca gregaria (desert locust)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
C;Accession: S12359
R;Marshall, J.; Buckingham, S.D.; Shingai, R.; Lunt, G.G.; Goosey, M.W.; Darlison, M.G.;
EMBO J. 9, 4391-4398, 1990
```



```

Db      187 NKNFVLDSDYKSGTWDIIEVPAYLN-VYEGDSNHPETDITFYIIIRKRTFYVNLIL 245
QY      529 PCVLIASMAILGFTLPDSDGKLSLGVITLLSLTVFLNMVAETMPATS----- 576
Db      246 PTVLISFLCVLFYLPAAEAGEKVTLGISILLVWFLLSKILPPTSLVPLIAKYLLF 305
QY      577 -----DAVPLWIRIVFLCWLPLWILKSRPGRPLILEFPTT 611
Db      306 TFMNTVSIIVTVIINWNERGPRTHRMFMYIISIFLHYLPAPLFMKRPRKTRLRWMMEM 365
QY      612 PCSDTSSERKHQILSDVELKERS-----KSLLANVLDIDDDFRHNC---RPMTPGGTLPH 664
Db      366 PGMSMPAHPHPSYSGPAELPKHISAIGGKQSKWEMVLSDLHHPNCKINKKVNSSGGE-- 423
QY      665 NPAFYTVVYQCGDGSIGPIGSTMPPDAVTHHTCIKSTYELGLLILKEIRFITDQLRKD 724
Db      424 -----GLGD-----GCRSESSDSILLSPEASKATE-----AVEFTAEHLRNE 462
QY      725 DECNDIANDWKFAAMVVDRLCLIIFTWFAITATIAVILSAPHI 767
Db      463 DLYIQTRDQKYAMVIDRLQLYIFPIVTTAGIVGILMDAPHI 505

```

Search completed: May 7, 2004, 11:39:30
 Job time : 20.738 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:18:28 ; Search time 14.3803 Seconds
(without alignments)
2788.120 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKLTVDDELWLAV.....MPAILATIAVLSSAPHIIVS 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1182.5	29.2	502	1	ACH7 MOUSE
2	1179.5	29.2	502	1	ACH7 CHICK
3	1173.5	29.0	502	1	ACH7 RAT
4	1159.5	28.7	502	1	ACH7 HUMAN
5	1157	28.6	499	1	ACH7 BOVIN
6	1104	27.3	498	1	ACH7 CAEL
7	888	22.0	567	1	ACH1 DROME
8	886	21.9	516	1	ACH1 MANSE
9	873	21.6	557	1	ACH1 SCHGR
10	848.5	21.0	529	1	ACH2 HUMAN
11	845	20.9	503	1	ACH3 HUMAN
12	843.5	20.9	576	1	ACH2 DROME
13	840	20.8	521	1	ACH3 DROME
14	838	20.7	528	1	ACH2 CHICK
15	827.5	20.5	456	1	ACHA CHICK
16	827.5	20.5	496	1	ACH3 CHICK
17	827	20.5	511	1	ACH2 RAT
18	823	20.4	499	1	ACH3 RAT
19	820.5	20.3	495	1	ACH3 BOVIN
20	820.5	20.3	512	1	ACH3 CARAU
21	815.5	20.2	519	1	ACH4 DROME
22	805	19.9	498	1	ACHP HUMAN
23	803	19.9	494	1	ACH6 CHICK
24	801	19.8	457	1	ACHA BOVIN
25	799	19.8	457	1	ACH6 HUMAN
26	797	19.7	454	1	ACHA MOUSE
27	794.5	19.7	457	1	ACH2 XENLA
28	792.5	19.6	493	1	ACH6 RAT
29	791.5	19.6	538	1	ACH8 CAEL
30	789	19.5	457	1	ACHA RAT
31	789	19.5	470	1	ACHP CHICK
32	787	19.5	491	1	ACHN CHICK
33	787	19.5	629	1	ACH4 MOUSE

34	786.5	19.5	456	1	ACHA_BRARE	Q88880 brachydanio
35	785	19.4	461	1	ACHA_TORMA	P02711 torpedo mar
36	783	19.4	495	1	ACHP RAT	P12392 rattus norv
37	782.5	19.4	500	1	ACHN RAT	P12390 rattus norv
38	782	19.3	502	1	ACHN_HUMAN	P17787 homo sapien
39	780	19.3	482	1	ACHA_HUMAN	P02708 homo sapien
40	780	19.3	627	1	ACHA_HUMAN	P43681 homo sapien
41	779.5	19.3	622	1	ACH4_CHICK	P09482 gallus gall
42	779	19.3	461	1	ACHA_TORCA	P02710 torpedo cal
43	773.5	19.1	457	1	ACH1_XENLA	P22456 xenopus lae
44	772	19.1	464	1	ACHO RAT	P12391 rattus norv
45	769.5	19.0	459	1	ACHN_CARAU	P19370 carassius a

ALIGNMENTS

RESULT 1

ID	ACH7 MOUSE	STANDARD;	PRT;	502 AA.
AC	P49582;	1996 (Rel. 33, Created)		
DT	01-FEB-1996	(Rel. 33, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Neuronal acetylcholine receptor protein, alpha-7 chain precursor.			
GN	CHRNA7 OR ACRA7.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BAIB/c; TISSUE=Brain;			
RX	MEDLINE=95324936; PubMed=7601470;			
RA	Orr-Urtreger A., Seldin M.F., Baldini A., Beaudet A.L.;			
RT	"Cloning and mapping of the mouse alpha 7-neuronal nicotinic			
RT	acetylcholine receptor."			
RL	Genomics 26:399-402(1995).			
CC	-!- FUNCTION: After binding acetylcholine, the AChR responds by an			
CC	extensive change in conformation that affects all subunits and			
CC	leads to opening of an ion-conducting channel across the plasma			
CC	membrane.			
CC	-!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-			
CC	bungarotoxin. The structure is probably pentameric (By			
CC	similarity).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- SIMILARITY: Belongs to the ligand-gated ionic channel family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L37663;	AAC42053.1;	-	
DR	PIR; A57175;	A57175.		
DR	MGI; MG1:99779;	Chrna7.		
DR	InterPro; IPR006029;	Neu_chan memb.		
DR	InterPro; IPR006202;	Neur_chan LBD.		
DR	InterPro; IPR006201;	Neur_chan.		
DR	Pfam; PF02931;	Neur_chan LBD; 1.		
DR	Pfam; PF02932;	Neur_chan memb; 1.		
DR	PRINTS; PR00252;	NR10CHANNEL.		
DR	TIGRFAMS; TIGR00860;	LIC; 1.		
DR	PROSITE; PS00236;	NEUROTE_ION_CHANNEL; 1.		
DR	Postsynaptic membrane; ionic channel; Glycoprotein; Signal;			
KW	Transmembrane; Multigene family.			
FT	SIGNAL	1 22	BY SIMILARITY.	
FT	CHAIN	23 502	NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,	
FT			ALPHA-7 CHAIN.	
FT	DOMAIN	23 230	EXTRACELLULAR (POTENTIAL).	

FT TRANSMEM 231 255 POTENTIAL.
FT TRANSMEM 262 280 POTENTIAL.
FT TRANSMEM 296 317 POTENTIAL.
FT DOMAIN 318 469 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 470 490 POTENTIAL.
FT FTISULFID 150 164 BY SIMILARITY.
FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION
(BY SIMILARITY).
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 502 AA; 56631 MW; C9312E5226D120E3 CRC64;
Query Match 29.28; Score 1182.5; DB 1; Length 502;
Best Local Similarity 45.87; Pred. No. 5.1e-75;
Matches 239; Conservative 80; Mismatches 118; Indels 85; Gaps 10;

QY 297 IYLNLSA---KVCAGYHKKRLHLDLPYNTLPRVNLNEDPQLSFLGLTLMQIIDVDE 353
Db 8 IWLALAAALHVSLOGEFORLYKELVKNYNPLRPVANDSQPLTVYFVLSLLQIMDVDE 67
QY 354 KNOLLVTNWLKLEWDMNLRWNTSDYGVKDLRIPPHRIWKPDVLMYNSADEGFGDTYQ 413
Db 68 KNOVLATTNIWLQSWTDBHYLQWNNSEYPGVKNVRFDPGQIWKPDILLYNSADERFDATEH 127
QY 414 TNNVVRNNGSLVYPGPIFKSTCKIDITWPFDDORCEMKEGWSWTYDGFOLDILODETG 473
Db 128 TNNLVNASGHCQVLPPIGPIKSSCYIDVRWPFDDVQCKIKFGWSYSGNSLIDLMQMBE--- 184
QY 474 GDISSYVLNGEWELLGVPGRKRIYNYNCCPEYDITFAIIRRTTYFFNLIIPCVLI 533
Db 185 ADISSYTPNSEDWLMGIPGRKNEKFEYCKEPEYDVTYVYTMRRRTLYGLNLLIPCVLI 244
QY 534 ASWALLGFTLPPDGEKLSLGVITLLSLTVPLNMVAETMPATSDAVPI----- 581
Db 245 SALALLVFLPADSGEKISLITVLLSLTVPLNVAETMPATSDVSLIAQYFASMTIIV 304
QY 582 -----WLRIVFLCWLPLWRMSRPG-----RLILFEFTTP 612
Db 305 GLSVVTVIVLRYHHHPDGGKMPKWTIRILLNCWFLRMKRGEDKVRPACQHKPR- 363
QY 613 CSUTSRRKHQILSDVELKERSKSLANVLIDDDFRHNCRPMTPGGTFLPHNPAFYRTV 672
Db 364 CS-----LASVELSAGAG-----PPTSNGNLLY--IGRGL 392
QY 673 YGQ-----GDDGSI-GPIGSTRMPDVAHTHCYKSTSEYELGLILKEIRFTDGLRDXDE 726
Db 393 EGMHCAPTSDSGVVGRLACSPTHDEHLMHGTHPSDGDPLAKILEEVRYIANFRQDCE 452
QY 727 CNDIANDWKAAMVVDRLCLIIFTWFAILATIAVLLSAPHII 768
Db 453 SEVICSEWKAACVVDRLCLMAFSVFTIICITIGLMSAPNEV 494

RESULT 2

ACH7 CHICK
ID -ACH7 CHICK STANDARD; PRT; 502 AA.
AC P22770;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91097796; PubMed=1702646;
RA Couturier S., Bertrand D., Matter J.M., Hernandez M.C., Bertrand S.,
RA Millar N., Valera S., Barkas T., Ballivet M.;
RT "A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is
RT developmentally regulated and forms a homo-oligomeric channel blocked
RL by alpha-BTX." Neuron 5:847-856(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90315158; PubMed=2369519;
RA Schoepfer R., Conroy W.G., Whiting P., Gore M., Lindstroem J.;
RT "Brain alpha-bungarotoxin binding protein cDNAs and MABs reveal
RT subtypes of this branch of the ligand-gated ion channel gene
RT superfamily." Neuron 5:35-48(1990).
RN [3]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=White leghorn; TISSUE=Erythrocyte;
RX MEDLINE=93049204; PubMed=1425587;
RA Matter-Sadzinski L., Hernandez M.C., Rostocil T., Ballivet M.,
RA Matter J.M.;
RT "Neuronal specificity of the alpha 7 nicotinic acetylcholine receptor
RT promoter develops during morphogenesis of the central nervous
RT system." EMBO J. 11:4529-4538(1992).
RN [4]
RP SEQUENCE OF 24-47.
RC TISSUE=Brain;
RX MEDLINE=85270494; PubMed=3860855;
RA Conti-Tronconi B.M., Dunn S.M.J., Barnard E.A., Dolly J.O., Lai F.A.,
RA Ray N., Rafferty M.A.;
RT "Brain and muscle nicotinic acetylcholine receptors are different but
RT homologous proteins." Proc. Natl. Acad. Sci. U.S.A. 82:5208-5212(1985).
RN [5]
RP MUTAGENESIS OF LEU-270.
RX MEDLINE=92049732; PubMed=1719423;
RA Recah F., Bertrand D., Galzi J.-L., Devillers-Thiery A., Mulle C.,
RA Hussy N., Bertrand S., Ballivet M., Changeux J.-P.;
RT "Mutations in the channel domain alter desensitization of a neuronal
RT nicotinic receptor." Nature 353:846-849(1991).
RN [6]
RP MUTAGENESIS TO CONVERT ION SELECTIVITY FROM CATIONIC TO ANIONIC.
RX MEDLINE=93024917; PubMed=1383829;
RA Galzi J.-L., Devillers-Thiery A., Hussy N., Bertrand S.,
RA Changeux J.-P., Bertrand D.;
RT "Mutations in the channel domain of a neuronal nicotinic receptor
RT convert ion selectivity from cationic to anionic." Nature 359:500-505(1992).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-
CC bungarotoxin. The structure is probably pentameric (by
CC similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- DEVELOPMENTAL STAGE: Alpha-7 transcripts transiently accumulate
CC in the developing optic tectum between E5 and E16.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X52295; CAA36543.1; -
DR EMBL; X68246; CAA48317.1; -
DR EMBL; X68586; CAA48576.1; -
DR PIR; JN0113; JN0113.
DR PDB; 1KC4; 17-APR-02.

similarity).

-!- SUBCELLULAR LOCATION: Integral membrane protein.

-!- MASS SPECTROMETRY: MW=54157.68; METHOD=MALDI.

-!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X70297; CAA49778.1; -

EMBL; U40583; AAA83561.1; -

EMBL; U62436; AAB40114.1; -

EMBL; Y08420; CAA69697.1; -

EMBL; AF385585; AAK68111.1; -

EMBL; L25827; -; NOT ANNOTATED_CDS.

EMBL; Z23141; CAA80672.1; -

EMBL; AF332758; AAK19515.1; -

PIR; G02259; G02259.

PIR; I37185; ACU047.

Genew; HGNC:1960; CHRNA7.

MIM; 118511; -

GO; GO:0005892; C:nicotinic acetylcholine-gated receptor-chan. . .; TAS.

GO; GO:0015464; F:acetylcholine receptor activity; TAS.

GO; GO:0004889; F:nicotinic acetylcholine-activated cation-se. . .; TAS.

GO; GO:0000187; P:activation of MAPK; TAS.

GO; GO:0006810; P:transport; TAS.

InterPro; IPR006029; Neu_chan memb.

InterPro; IPR006202; Neu_chan IBD.

Pfam; PR02931; Neu_chan_IBD; 1.

Pfam; PR02932; Neu_chan_memb; 1.

PRINTS; PR00252; NRIONCHANNEL.

TIGRFAMS; TIGR00860; LIC; 1.

PROSITE; PS00236; NEUROTR ION CHANNEL; 1.

Postsynaptic membrane; Ionic Channel; Glycoprotein; Signal; Transmembrane; Multigene family.

FT SIGNAL 23 22

FT CHAIN 1 502

FT DOMAIN 23 230

FT TRANSMEM 231 255

FT TRANSMEM 262 280

FT TRANSMEM 296 317

FT DOMAIN 318 469

FT TRANSMEM 470 490

FT DISULFID 150 164

FT DISULFID 212 213

FT CARBOHYD 46 46

FT CARBOHYD 90 90

FT CARBOHYD 133 133

FT CONFLICT 11 11

FT CONFLICT 58 58

FT CONFLICT 134 134

FT CONFLICT 364 364

FT CONFLICT 375 375

FT CONFLICT 409 413

SQ SEQUENCE 502 AA; 56449 MW; D94B3A482EAA0E42 CRC64;

Query Match 28.7%; Score 1159.5; DB 1; Length 502;

Best Local Similarity 45.1%; Pred. No. 2e-73;

Matches 237; Conservative 82; Mismatches 115; Indels 91; Gaps 12;

QY 297 IYINLSAK---VCLAGYHEKRLHLLDPNTLERPVNESDPLQLSEGLTMOIIDVDE 353

DB 8 VMLAALASLLHVSLOQEFGRKLYKELVKNVPLERPVANDSQLTIFYSLSLQIMDVDE 67

QY 354 KNQLLVNTWVLEWMDMLRNWNTSDYGGVKDLRIPPHRIWKPDLVLMYNSADEGFGDTYQ 413

106 68 KNQVLITINWLMQSWTDHYLQNNVSEYPGVKTYRFPDGIQWIKPDILLYNSADERPDATFH 127

QY 414 TNVVNRNGSLYVPPGIFKSTCKIDITWPFDDQRCMFKSGWTYDGFOLDLQDETGT 473

DB 128 TNVLVNSSGHCQYLPPIGIFKSSCYIDVRWFFDVQHCCKLFGWSYGGWSLDLQMQE- 184

QY 474 GDTSYVVLNGEWELLGVGRNEIYVNCCEPEYDITFAIIIRRTLYYFNLIIPCVLI 533

DB 185 ADISGYPNGEWDLVIGPKRSERFYECCKPEYDVTFTVTRRTLYYGLNLLIPCVLI 244

QY 534 ASMALLGFTLPDPSGKSLGLVTILLSTVFLNVAETMPATSDAVPL----- 581

DB 245 SALALLVFLPADSGEKISLGIIVLISLTVFMLVAEIMPATSDSVPLIAQYFASWTIIV 304

QY 582 -----WTRIVFLCWLPLMRLMSRPG-----RPLILEPPTTP 612

DB 305 GLSVVTVTVILQYHHHDPDGGKMPKTRVILLNWCWFLRMKRPGEKVRP-ACQHKQR 363

QY 613 CSDTSSERKHQILSDVELKERSKSLANVLDDDDFRH-----NCRPMTGGTLPHPNPAF 668

DB 364 CSLASVE-----MSAVAPPASNGNLLY-----IGFRGLDGVHCVF-TP----- 401

QY 669 YRTVYGGDGSIGPIGSTRMPDAVTH-----HTCIKSTSEYELGLIKIRFITDOLRK 723

DB 402 -----DSGVV-----CGRMACSPHDELLHGGQPPGPDPLAKILEVRYIANRFR 449

QY 724 DDECNDIANDKPAAMVVRDLCLIIITMFAITATIAVLLSAPHII 768

DB 450 QDESEAVCSEWKAACVVDRLCLIMAFSVFTICTIGILMSAPNFV 494

RESULT 5

ACH7_BOVIN STANDARD; PRT; 499 AA.

AC P54131;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.

GN CHRNA7.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).

RC TISSUE=Adrenal medulla;

RX MEDLINE=95346009; PubMed=7620615;

RA Garcia-Guzman M.; Sala F.; Sala S.; Campos-Caro A.; Stuehmer W.; Gutierrez L.; Criado M.;

RT "Alpha-Bungarotoxin-sensitive nicotinic receptors on bovine chromaffin cells: molecular cloning, functional expression and alternative splicing of the alpha 7 subunit.";

RL Eur. J. Neurosci. 7:647-655(1995).

CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an extensive change in conformation that affects all subunits and leads to opening of an ion-conducting channel across the plasma membrane.

CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-bungarotoxin. The structure is probably pentameric (By similarity). Homo-oligomer of the short form gives rise to unfunctional channels, as does coexpression of both long and short forms of the receptor.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=2;

CC Name=Long;

CC isoId=P54131-1; Sequence=Displayed;

CC Name=Short;

CC isoId=P54131-2; Sequence=VSP 000075;

CC -!- TISSUE SPECIFICITY: At least in chromaffin cells.

CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X93604; CAA63802.1; --
 CC InterPro; IPR006029; Neu channel memb.
 CC InterPro; IPR006202; Neur_chan_LBD.
 CC InterPro; IPR006201; Neur_chan_LBD.
 CC Pfam; PF02931; Neur_chan_LBD; 1.
 CC Pfam; PF02932; Neur_chan_memb; 1.
 CC PRINTS; PR00252; NRIONCHANNEL.
 CC TIGRFAMS; TIGR00860; LIC; 1.
 CC PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 CC Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 CC Transmembrane; Multigene family; Alternative splicing.
 KW SIGNAL 1
 FT CHAIN 20 499
 FT
 FT DOMAIN 20 227
 FT TRANSMEM 228 252
 FT TRANSMEM 259 277
 FT TRANSMEM 293 314
 FT TRANSMEM 315 466
 FT TRANSMEM 467 487
 FT DISULFID 147 161
 FT DISULFID 209 210
 FT
 FT CARBOHYD 43 43
 FT CARBOHYD 87 87
 FT CARBOHYD 130 130
 FT VARSPIC 262 290
 FT
 FT SEQUENCE 499 AA; 56002 MW; ABE5D0B3820D42D5 CRC64;
 Query Match 28.6%; Score 1157; DB 1; Length 499;
 Best Local Similarity 46.4%; Pred. No. 3e-73;
 Matches 235; Conservative 76; Mismatches 122; Indels 74; Gaps 10;
 QY 305 VCLAGYHEKRLHLLDPYNTLEPVLNESPLOLSCGLTLMQIDVDKQMLVTVNWL 364
 DB 16 VSLQGEQRKLYKDLVKNYNLEPVRVANDSLPVFVLSLQIMVDKQVLTITNWL 75
 QY 365 KLEWDMNLRNNTSDYGVGKDLRIPPHRIKWPDKVLMVNSADEGFGTYQTIVVVRNNGSC 424
 DB 76 QMTWTDHYLQWNASEYPGKTVRFPDGGIWKPDILLYNSADERFDTAFTNVLVNSGHC 135
 QY 425 LYVPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGFOLDLOQDETGGDISVVLNGE 484
 DB 136 QYLPFGIFKSSCYIDVRFPFPDVOQCKLKFGSWYGGWSLDLQWQE--ADISGYIPNGE 192
 QY 485 WELGVGPKNEIYVNCPEYIDITFAIIRRTLYFFNLIIPCVLIASMAILGFTLP 544
 DB 193 WDLVGVLGKSEKEFYECKEYPDVTFTVSTRRTLYGLNLLIPCVLISALALLVFLP 252
 QY 545 PDSGEKLSGLVTLLSLTVFLNMAVETMPATSDAVPL----- 581
 DB 253 ADSGEKLSGLVTLLSLTVFLNMAVETMPATSDAVPLIAQYFASMTIIVGLSVVTVIVL 312
 QY 582 -----WIRIVFLCWLPWILRMSPGRPLILEFTPTPCSTSSERKHQIUSD 627
 DB 313 QYHHDDPGGKPKWTRVLLNWCWFLRMKRP-----EDKVRPACQNNRRCS--LAS 365
 QY 628 VELKE-RGSKSLANVLDDIDDFR-----HNCRPMTPPGGTLPHNPAFYRTVYGGQDGGSI- 681
 DB 366 VEMSAVAGPPATNGLLYI--GFRGLDTHCAP-TP-----DSGVVC 404
 QY 682 GPIGSTRMPDVAHTHTCKTSYTEYELGLILKEIRPITQLRKDDCNDIANDKFPAMV 741

DB 405 GRVACSPTHDEHLLHAGQSPGPDLLAKILEVRYIAHRFCQDESAVCSWKFAACVV 464
 QY 742 DRCLLIIFTMFAITAVLLSAPHII 768
 DB 465 DRCLMAFVSFTILCTIGILMSAPNFV 491
 RESULT 6
 ACHEL CAEEL
 ID ACHEL CAEEL STANDARD; PRT; 498 AA.
 AC P48180;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acetylcholine receptor like protein, alpha-type chain precursor.
 GN F25G6.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=36196478; PubMed=8627624;
 RA Ballivet M., Alliod C., Bertrand S., Bertrand D.;
 RT "Nicotinic acetylcholine receptors in the nematode Caenorhabditis
 RT elegans".
 RL J. Mol. Biol. 258:261-269 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Nelson J., Wohldmann P.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Possible acetylcholine receptor.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X83887; CAA58764.1; --
 CC EMBL; AF022973; AAC25796.1; --
 CC PIR; S68588; S68588.
 CC HSP; P58154; 119B.
 CC WormPep; F25G6.3; CB09639.
 CC InterPro; IPR006029; Neu channel memb.
 CC InterPro; IPR006202; Neur_chan_LBD.
 CC InterPro; IPR006201; Neur_chan_LBD.
 CC Pfam; PF02931; Neur_chan_LBD; 1.
 CC Pfam; PF02932; Neur_chan_memb; 1.
 CC PRINTS; TIGR00860; LIC; 1.
 CC TIGRFAMS; TIGR00860; LIC; 1.
 CC PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 CC Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein;
 KW Transmembrane; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 498
 FT
 FT DOMAIN 20 230
 FT TRANSMEM 231 252
 FT TRANSMEM 261 279
 FT TRANSMEM 295 314
 FT TRANSMEM 315 472
 FT TRANSMEM 473 493
 FT DISULFID 147 161
 FT DISULFID 211 212
 FT
 FT CARBOHYD 43 43
 FT
 FT ACETYLCHOLINE RECEPTOR LIKE PROTEIN,
 FT ALPHA-TYPE CHAIN.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT POTENTIAL.
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT POTENTIAL.
 FT BY SIMILARITY.
 FT ASSOCIATED WITH RECEPTOR ACTIVATION
 FT (BY SIMILARITY).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
 GN CHRNA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Hypothalamus;
 RX MEDLINE=97062879; PubMed=8906617;
 RA Elliott K.J., Ellis S.B., Bercikan K.J., Urrutia A.,
 RA Chavez-Noriega L.E., Johnson B.C., Velicelebi G., Harpold M.M.;
 RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta
 RT 2-beta 4 nicotinic acetylcholine receptor subunits and functional
 RT expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and
 RT beta 4 subunits.";
 RL J. Mol. Neurosci. 7:217-228(1996).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP Groot Kormelink P.J.;
 RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Blechschmidt K., Rosenthal A.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 CC extensive change in conformation that affects all subunits and
 CC leads to opening of an ion-conducting channel across the plasma
 CC membrane.
 CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different types
 CC of subunits: alpha and non-alpha (beta). Alpha-2 subunit can be
 CC combined to beta-2 or beta-4 to give rise to functional receptors.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U62431; AB040109.1; -;
 DR EMBL; Y16281; CA076154.1; -;
 DR EMBL; AF311103; -; NOT ANNOTATED_CDS.
 DR Genew; HGNC:1956; CHRNA2.
 DR MIM; 118502; -;
 DR GO; GO:0005892; C:nicotinic acetylcholine-gated receptor-chan. . .; TAS.
 DR GO; GO:0013464; F:fractylcholine receptor activity; TAS.
 DR GO; GO:0004889; F:nicotinic acetylcholine-activated cation-se. . .; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR GO; GO:0007268; P:synaptic transmission; TAS.
 DR InterPro; IPR006029; Neu_chan_LBD.
 DR InterPro; IPR006202; Neu_chan_LBD.
 DR InterPro; IPR006201; Neu_chan_LBD.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_LBD; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMs; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 26 POTENTIAL,
 FT CHAIN 27 529 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 FT ALPHA-2 CHAIN.
 FT DOMAIN 27 264 EXTRACELLULAR.
 FT TRANSMEM 265 289 POTENTIAL.
 FT TRANSMEM 297 315 POTENTIAL.
 FT TRANSMEM 331 352 POTENTIAL.
 FT DOMAIN 353 502 CYTOPLASMIC.

FT TRANSMEM 503 521 POTENTIAL.
 FT DISULFID 183 197 BY SIMILARITY.
 FT DISULFID 247 248 ASSOCIATED WITH RECEPTOR ACTIVATION
 FT (BY SIMILARITY).
 FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 125 125 A -> T (IN REF. 3).
 SQ SEQUENCE 529 AA; 59735 MW; 7F512B06CCD9AAFD CRC64;
 Query Match 21.0%; Score 848.5; DB 1; Length 529;
 Best Local Similarity 36.7%; Pred No. 1e-51;
 Matches 183; Conservative 87; Mismatches 154; Indels 75; Gaps 10;
 QY 309 GYH---EKRLIHDLDPYNTLERPVNESDPLQSLFGLTLMQIIDVDEKQOLLVTNVLWK 365
 Db 53 GSHTEDEDLFKLFRGYNRWARPVNTSDVIVRFGLSIAQLIDVDEKQMTTNVLWK 112
 QY 366 LEWDMMLRWNTSDYGVKDLRIPPHIRKPDVLMYNSADEGDTGTNTVVVNRNNGSCL 425
 Db 113 QEWSDYKLRWNPADFGNITSLRVSEMIWIPDIVLYNNADEGFAVHTMTKAHLFSTGVH 172
 QY 426 YVPPGIEPKSTCKIDITWPPDDQRCMKFGSWTYDGFQDLQLQDETGGDISSYVLNGEW 485
 Db 173 WPPAIYKSSCSIDVTFPPDQONCKMKFGSWTYDKAKIDLEQMEQT-VDLKDYWESGEW 231
 QY 486 ELLGVGPKRNEIYVNCPEYIDITFAIIRRTLYVFFNLIIPCVLIASVALLGFTLPP 545
 Db 232 AIVNATGYNKSKYDCCAEIYPDVYAFVIRRLPLFVTINLIIPCLISCLTIVLVFYLPS 291
 QY 546 DSGKLSLGVILLISLTVFLNMVAETMPATSDAVPL----- 581
 Db 292 DCGEKITLCISVLLSLTVFLLLITETIPSTSLVPLIGEYLLFTMTFVLSIVITVFLVN 351
 QY 582 -----WIRVFLCWLPMILRMSPGRPLILEFTPTTCSDTSSRKHQILSDV 628
 Db 352 VHRSPSTHTPHWVRGALLCCVPRWLLMNPPEPVEL---CHPLRLKLSPSYHLESNV 408
 QY 629 ELKERSKSLANVLIDDDDFHNCR-PMTPG-GTLPHPNAPYRTVYGGDGGSGTPTGS 686
 Db 409 DAERE-----VVVEEDRWACAGHVASVGT-----CSGHLHSGASGPKAE 452
 QY 687 TRMPDA---VTHHTCIKSTVEYELGLILKEIRFIDQLRKDDKDCNDIANDKFAAMVDR 743
 Db 453 ALLQGEILLSPH-----MOKALEGVHYIADHLRSEDADSVKEDWKYVAVVDR 502
 QY 744 LCLIFTWFAILLATIAVLL 762
 Db 503 IFLWLFIVCFLTGTLGLFL 521

RESULT 11

ACH3 HUMAN
 ID ACH3 HUMAN STANDARD; PRT; 503 AA.
 AC P32977; Q15823; Q96RH3; Q99553; Q9BQ93;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-3 chain precursor.
 GN CHRNA3 OR NACHRA3
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=90245296; PubMed=2336208;
 RA Fornasari D., Chini B., Tarroni P., Clementi F.;
 RT "Molecular cloning of human neuronal nicotinic receptor alpha
 RT 3-subunit";
 RL Neurosci. Lett. 111:351-356(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Thymus;
 RX MEDLINE=91114756; PubMed=1989896;
 RA Mihovilovic M., Roses A.D.;
 RT "Expression of mRNAs in human thymus coding for the alpha 3 subunit
 of a neuronal acetylcholine receptor.";
 RL Exp. Neurol. 111:175-180(1991).
 [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97062879; PubMed=8906617;
 RA Elliott K.J., Ellis S.B., Berckhan K.J., Urrutia A.,
 Chavez-Noriega L.E., Johnson E.C., Velicelebi G., Harpold M.M.;
 RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta
 2-beta 4 nicotinic acetylcholine receptor subunits and functional
 expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and
 beta 4 subunits.";
 RL J. Mol. Neurosci. 7:217-228(1996).
 [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=97162233; PubMed=9009220;
 RA Groot Kormelink P.J., Luyten W.H.M.L.;
 RT "Cloning and sequence of full-length cDNAs encoding the human neuronal
 nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and
 expression of seven nAChR subunits in the human neuroblastoma cell
 line SH-SY5y and/or IMR-32.";
 RL FEBS Lett. 400:309-314(1997).
 [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=99118870; PubMed=9921897;
 RA Rempel N., Heyers S., Engels H., Slegers E., Steinlein O.K.;
 RT "The structures of the human neuronal nicotinic acetylcholine receptor
 beta2- and alpha3-subunit genes (CHRNA2 and CHRNA3).";
 RL Hum. Genet. 103:645-653(1998).
 [6]
 RP SEQUENCE FROM N.A. AND VARIANT LEU-21 INS.
 RX MEDLINE=21342809; PubMed=11450844;
 RA Lev-Lehman E., Bercovich D., Xu W., Stockton D.W., Beaudet A.L.;
 RT "Characterization of the human beta4 nAChR gene and polymorphisms in
 CHRNA3 and CHRNA4.";
 RL J. Hum. Genet. 46:362-366(2001).
 [7]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [8]
 RP SEQUENCE OF 30-503 FROM N.A.
 RX TISSUE=Brain;
 RA Anand R., Lindstrom J.;
 RT Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.
 [9]
 RP SEQUENCE OF 6-493 FROM N.A.
 RX TISSUE=Epidermal keratinocytes;
 RA Arredondo J., Grando S.A.;
 RT "Cloning cholinergic receptors in human keratinocytes.";

Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 extensive change in conformation that affects all subunits and
 leads to opening of an ion-conducting channel across the plasma
 membrane.
 -!- SUBUNIT: Neuronal AChR seems to be composed of two different type
 of subunits: alpha and non-alpha (beta).
 -!- SUBCELLULAR LOCATION: Integral membrane protein.
 -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Name=1;
 IsoId=P32297-1; Sequence=Displayed;
 Name=2;
 IsoId=P32297-2; Sequence=VSP_000073;
 Note=No experimental confirmation available;
 -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).

 EMBL; M86383; AAC84176.1; -;
 EMBL; M37981; AAA59942.1; -;
 EMBL; U62432; AAB40110.1; -;
 EMBL; Y08418; CAA69695.1; -;
 EMBL; AJ007783; CAA07682.1; -;
 EMBL; AJ007784; CAA07682.1; JOINED.
 EMBL; AJ007785; CAA07682.1; JOINED.
 EMBL; AJ007786; CAA07682.1; JOINED.
 EMBL; AJ007787; CAA07682.1; JOINED.
 EMBL; BC001642; AAH01642.1; -;
 EMBL; BC002996; AAH02996.1; -;
 EMBL; BC000513; AAH00513.1; -;
 EMBL; AF385584; AAH68110.1; -;
 EMBL; X53559; CAA37625.1; -;
 PIR; A37040; A37040.
 PIR; A53956; A53956.
 Genew; HGNC:1957; CHRNA3.
 MIM; 118503; -;
 GO; GO:0005892; Nicotinic acetylcholine-gated receptor-chan. . . ; TAS.
 GO; GO:0015464; Facetylcholine receptor activity; TAS.
 GO; GO:0004889; Fnicotinic acetylcholine-activated cation-se. . . ; TAS.
 GO; GO:0005215; F:transporter activity; TAS.
 GO; GO:0007165; P:signal transduction; TAS.
 GO; GO:0006810; P:transport; TAS.
 InterPro; IPR006029; Neu channel memb.
 InterPro; IPR006202; Neur chan LBD.
 InterPro; IPR006201; Neur channel.
 Pfam; PF02931; Neur chan LBD; 1.
 Pfam; PF02932; Neur chan memb; 1.
 PRINTS; PR00252; NRIONCHANNEL.
 TIGRfams; TIGR00860; LIC; 1.
 PROSITE; PS00236; NEUOTR ION CHANNEL; 1.
 Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 Transmembrane; Multigene family; Alternative splicing; Polymorphism.
 SIGNAL 1 29 POTENTIAL.
 CHAIN 30 503
 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 ALPHA-3 CHAIN.
 EXTRACELLULAR (POTENTIAL).
 DOMAIN 30 238
 TRANSMEM 239 263 POTENTIAL.
 TRANSMEM 271 289 POTENTIAL.
 TRANSMEM 305 326 POTENTIAL.
 DOMAIN 327 475 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 476 495 POTENTIAL.
 DISULFID 157 171 BY SIMILARITY.
 DISULFID 221 222 ASSOCIATED WITH RECEPTOR ACTIVATION
 (BY SIMILARITY).
 CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT VARSPLIC 1 5 MALAV -> MGSGLP (in isoform 2).
FT FT /FtId=VSP_000073.
FT VARIANT 21 21 L -> LL.
FT /FtId=VAR_013240.
FT CONFLICT 5 13 VSLPLALSP -> ALAAPGAVA (IN REF. 2).
FT CONFLICT 11 14 LSPP -> CRA (IN REF. 1).
FT CONFLICT 100 100 D -> G (IN REF. 1).
FT CONFLICT 132 133 DD -> TT (IN REF. 1).
FT CONFLICT 235 235 I -> S (IN REF. 1).
FT CONFLICT 430 430 L -> V (IN REF. 1).
SQ SEQUENCE 503 AA; 57309 MW; 8A9EBC5D71AEC7D6 CRC64;

Query Match 20.9%; Score 845; DB 1; Length 503;
Best Local Similarity 35.2%; Pred. No. 1.7e-51;
Matches 178; Conservative 97; Mismatches 166; Indels 64; Gaps 6;

QY 295 LLIYINLSAKYCLAGYHEKRLHLLDLPYNTLERPVLESPLQSLFGLTLMQIIDVDEK 354
D 17 LLLLSL-LPVARASEAHLRFLERLFEDYNEIIRPVANVSDFVTHFEVMSQLVKVDEV 75

355 NOLLVTNVLKLEWDMNLRNWTSDYGVKDLRIPPHRIWKPDVLMYNSADEGFGTYQT 414
D 76 NQIMETNLWLQIWNIDYKLNWPSDYGAEFMRVPAQKIWKPDIVLYNNVAGDFQVDDKT 135

415 NVVVRNNGSLYVPPGIFKCKIDITWPPDDORCEMKFGSWITVDFGQLDQLQDETGG 474
D 136 KALLKYTGVTWIPPAIFKSKKIDVTFFDYQNTWKFGSWYDKAKIDLVLIG-SSM 194

475 DISSVYNGEWELLGVPGKRNIYVNCPEYIDITFAIIRRTLYYFFNLIIPCVLIA 534
D 195 NLKDYWESGEWAIIKAPGYKHIDYKNCCEIYPDITYSLYIRRLPLFLYTNLIIPCLLIS 254

535 SMALLGFTLPDPSGKLSIGVTLLISLVFLNVAETWPAVPL----- 581
D 255 FLTVLVFLPSDCGKVTLCISVLSLTVFLLVITITPSTLSVLPLIGVLLFMFIW 314

582 -----WIRVFLCWLFWILMRSPGRPLILEFPTPCSDTS 617
D 315 LSIVITVFLNVHYRTPTTHMPSWKVFLNLLPRV-----FMTRTPSNEG 362

618 SERKHQILSDVELKERSKSLIANVLIDDDFRHNCRPMTPGCTLPHNPAFVRTVYGGD 677
D 363 NAQPRPLRYGAEISLNCFSRAES-----KGCKEGYPCQDGMCGYCHRRIKISNF 413

678 DSGSIGPSTRMPDVAHTHTCTKSSTYELGLILKEIRFTDOLRAKDCNDIANDWKA 737
D 414 SANLTRSSSSVDVAVLS-----LSALSPETKEATIOSVKYIAENKAQNEAKIQQDWKV 469

738 AMVVDRLCLIIFTFAILATIAVLL 762
D 470 AMVIDRIFLWFLVLCILGTAGLFL 494

RESULT 12
ID ACH2 DROME STANDARD; PRT; 576 AA.
AC P17644; Q9VC73;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Acetylcholine receptor protein, alpha-like chain 2 precursor.
GN NACR-ALPHA-96AB OR ACRE OR SAD OR ACR96AB OR CG6844.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RX Baumann A., Jonas P., Gundelfinger E.D.;
RA "Sequence of D alpha 2, a novel alpha-like subunit of Drosophila
```

```
RT nicotinic acetylcholine receptors.";
RL Nucleic Acids Res. 18:3640-3640(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RX MEDLINE=90353591; PubMed=2117557;
RA Jonas P., Baumann A., Merz B., Gundelfinger E.D.;
RA "Structure and developmental expression of the D alpha 2 gene
RT encoding a novel nicotinic acetylcholine receptor protein of
RT Drosophila melanogaster."
RL FEBS Lett. 269:261-268(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90360975; PubMed=1697262;
RX Sawruk E., Schloss P., Betz H., Schmitt B.;
RT "Heterogeneity of Drosophila nicotinic acetylcholine receptors: SAD,
RT a novel developmentally regulated alpha-subunit."
RL EMBO J. 9:2671-2677(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cuiway S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Head;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
```

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: CNS in embryos.
 CC -!- DEVELOPMENTAL STAGE: Late embryonic and late pupal stages.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X52274; CAA36517.1; -;
 DR EMBL; X53583; CAA37652.1; -;
 DR EMBL; AE003748; AAF56303.1; -;
 DR EMBL; AY058446; AAL13675.1; -;
 DR PIR; S11679; ACFFA2.
 DR FlyBase; FBgn0000039; nAcR-alpha-96Ab.
 DR InterPro; IPR006029; Neu channel memb.
 DR InterPro; IPR006202; Neur chan LBD.
 DR InterPro; IPR006201; Neur channel.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMs; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUOTR ION CHANNEL; 1.
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 21 PROBABLE.
 FT CHAIN 22 576 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-
 FT LIKE CHAIN 2.
 FT DOMAIN 22 261 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 262 285 POTENTIAL.
 FT TRANSMEM 293 311 POTENTIAL.
 FT TRANSMEM 327 346 POTENTIAL.
 FT DOMAIN 347 526 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 527 545 POTENTIAL.
 FT DISULFID 169 183 BY SIMILARITY.
 FT DISULFID 243 244 ASSOCIATED WITH RECEPTOR ACTIVATION
 FT (BY SIMILARITY).
 FT CARBOHYD 65 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 254 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 570 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 576 AA; 65506 MW; 97D6A46CADCF42F CRC64;
 Query Match 20.9%; Score 843.5; DB 1; Length 576;
 Best Local Similarity 34.4%; Pred. No. 2.6e-51;
 Matches 193; Conservative 87; Mismatches 170; Indels 111; Gaps 11;
 QY 290 HSW-----IFLLIVNLSAKVLGAYHEKRLHDLDPVNTLRPPVLNESDPLQLSGLT 344
 DB 18 HIWCKPCLLLVLLLCETVQANPPAKLYDLDLSYNNRLRPVSNNTDTLVKLGLR 77
 QY 345 LMQIIDVDENQLLVNVLKLEWNNMLRWNTSDYGVKDLRTPPHRWKPVLMVNSA 404
 DB 78 LSQILDLNLKQDILTITNVLEHEWQDKFKWDFSEYGVTELYVPSHIWLPDIVLYNA 137
 QY 405 DEGFDTGYQTNVVRNNGSLYVPPGIFKSTCKIDITWFPDQRCMKFGSWTYDGFOL 464
 DB 138 DGEYVVTMTKALHYTGKVVVTPPAIFKSSCEIDVRYFPDQTCFMKFGSWTYDQDI 197
 QY 465 DL-----QLQD-----ETGGDISSVLNGEWELLVGPKNEIYVNCPEYIDITAI 515
 DB 198 DLKHSQNDKXKVEIGIDREYVPSVEMDIIGVPAERHEKYPCCAEYPIDFNITL 257
 QY 516 RRTLYFFNLLIPCVLIASMLLGLTLPDPSGKLSGLVTLLSLTFLNLMVAETMPAT 575
 DB 258 RRKTLFTVNLIIPCVGISVLSVLFVLPADSEKALCISILLSQTMFLLLISEIPST 317
 QY 576 SDAVPL-----WIRVFLCWLPWIRMS 598
 DB 318 STALPLLGKYLFTMLLVGLSVVITIIILNIHVKPSTHQMRWIRSFIFKRLPKLLMR 377

QY 599 RPRGLIILEFPTTPCSDTSSERKHQILSDVELKE-----RSSKSLIANVLDDDDFRHNC 653
 DB 378 VP-----KDLLRLAANKINYGKSKTFEGQAL-MDE----- 409
 QY 654 RPTMPGGTLPHPAPRYTYVQGD-----DGSIGPIGS-----TRMPDA 692
 DB 410 MQMNSGGSPDLRLMRQGRVAGCGNMGHVTTATNFRSLGVLGALGGSLSTLSGYNGLPV 469
 QY 693 VT-----HHTCIKSSTVEYELGLILKEIRFTDQLKDECDNDIANDKFAAMVVDRLCL 746
 DB 470 LSGLLDLSLSDVAARKKYPFELEKAIHNVMPFIQHEMQRODEFNAEDQDQMGFVAMVMDRLFL 529
 QY 747 IIFTMFAILATIALLSAPHI 767
 DB 530 WLFMTIASLVGTFTVILGEAPSL 550
 RESULT 13
 ID ACH3_DROME STANDARD; PRT; 521 AA.
 AC P04755; Q9VZC3;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Acetylcholine receptor protein, beta-like chain 1 precursor.
 GN NACR-BETA-64B OR ACRD OR ARD OR ACR64B OR CG11348/CG12606.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_taxid=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Hermans-Borgmeyer I., Zopf D., Ryseck R.-P., Hovenann B., Betz H.,
 RA Gundelfinger E.D.;
 RA "Primary structure of a developmentally regulated nicotinic
 RT acetylcholine receptor protein from Drosophila.";
 RL EMBO J. 5:1503-1508(1986).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=88298842; PubMed=3136037;
 RA Sawruk E., Hermans-Borgmeyer I., Betz H., Gundelfinger E.D.;
 RT "Characterization of an invertebrate nicotinic acetylcholine receptor
 RL gene: the ard gene of Drosophila melanogaster.";
 RN FEBS Lett. 235:40-46(1988).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=88174720; PubMed=2832736;
 RA Wadsworth S.C., Rosenthal L.S., Kammermeyer K.L., Potter M.B.,
 RA Nelson D.J.;
 RT "Expression of a Drosophila melanogaster acetylcholine receptor-
 RL related gene in the central nervous system.";
 RN Mol. Cell. Biol. 8:778-785(1988).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.H., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: CNS in embryos.
CC -!- DEVELOPMENTAL STAGE: Late embryonic and late pupal stages.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04016; CAA27641.1; -;
CC EMBL; X07956; CAA30778.1; -;
CC EMBL; X07957; CAA30778.1; JOINED.
CC EMBL; X07958; CAA30778.1; JOINED.
CC EMBL; M20316; AAA28311.1; -;
CC EMBL; AE003481; AAF47900.1; -;
CC PIR; S03012; ACFFNN.
CC FlyBase; FBgn0000038; nAChR-beta-64B.
CC InterPro; IPR006029; Neu channel memb.
CC InterPro; IPR006202; Neur_chan_LBD.
CC InterPro; IPR006201; Neur_chan_LBD.
CC Pfam; PF02931; Neur_chan_LBD; 1.
CC Pfam; PF02932; Neur_chan_memb; 1.
CC PRINTS; PR00252; NRIONCHANNEL.
CC TIGRFAMs; TIGR00860; LIC; 1.
CC PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
CC Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
CC Transmembrane; Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 521 ACETYLCHOLINE RECEPTOR PROTEIN, BETA-LIKE
FT CHAIN 1.
FT FT
FT DOMAIN 25 235
FT TRANSMEM 236 260
FT TRANSMEM 268 286
FT TRANSMEM 302 323
FT DOMAIN 324 481
FT TRANSMEM 482 500
FT DISULFID 152 166
FT CARBOHYD 48 73
FT VARIANT 73 73
FT CONFLICT 383 384 EL -> DV (IN REF. 3).
FT SEQUENCE 521 AA; 59901 MW; FF9BA2ABC0C3AA62 CRC64;
SQ

Query Match 20.8%; Score 840; DB 1; Length 521;
Best Local Similarity 35.6%; Pred. No. 3.9e-51;
Matches 186; Conservative 95; Mismatches 172; Indels 70; Gaps 9;
QY 291 SWIFLLIYMLSAKVLACGLAGHKKLLHLLDPYNTLPRPVNLSEDPQLSLGFLMLQIID 350
DB 7 SWLLCSLVLVAFSLVSASEDEERLVRDLFRGVNKLIRPVQNMTPKQVGVRRGLAFVQLIN 66
QY 351 VDEKNQLLVNMLKLEWNNMLRWNTSDYGVVNDLIRPPHRIWKPDLVMTNSADEGPDG 410
DB 67 VNEKNQVKNVNLRLVWYQYQLOWDADYGGIGLRLPDKPKWKPDVILFNNADGNVEV 126
QY 411 TYQNVVVRNNGSLYVPPGIFKSTKIDITWPPFDQRCMKFGSTYDGFQDLQL-Q 469
DB 127 RYKSNVLIYPTGEVLWVPPAIYQSSCTIDVTYFPDQDTCIMKFGSWTFNGDQVSLALYN 186
QY 470 DETGGDLSVVNGEWELLGVGPKRNEIYNCCPEP-YIDITPAIIIRRRRLTYFFENLI 528
DB 187 NKNFVDSLSDYVKSQGTWIIIEVPAVLN-VYEGDSNHPETDITFYIIIRKTLFTVNLIL 245
QY 529 PCVLIASMALIGFTLPDPSGKLSGLVTILLSLTVFLNMVAETMPATS----- 576
DB 246 PTVLISFLCVLFLPAEAGEKVTIGISILLSLVFLVLLVSKILLPPTSLVLIKAYLLF 305
QY 577 -----DAVPLMIRIVFLCWLPIWILMRSPGRLILEPPTT 611
DB 306 TFIIMTVSILVTVIIINWFRGPRTRMPMYIRSGIFLHYLPAFLFMKPRKTRLRMMEM 365
QY 612 PCSTSSERKHQILSDVELKERS-----KSLANVLDDDDERHNC---RPMTPGGTLPH 664
DB 366 PGMSMPAHPHPSYSPALPKHISAIGKQSKQMEVMSLSDHHPNCKINRKVNSGGEJ-- 423
QY 665 NPAFVRTVYGGDDSGISGIPGSTMPDAVTHHTCIKSTYELGLILKEIRFITDQLRDK 724
DB 424 -----GLGD-----GCRRESSDSILLSPEASKATE-----AVEFIAEHLRNE 462
QY 725 DECNDIANDKPAAMVVDRLCLIIPTMPAILATIAVLSAPHI 767
DB 463 DLYIOTREDWKYAMVDRQLQIYFIVTGTAGTIGILMDAPHI 505

RESULT 14
ACH2 CHICK
ID ACH2 CHICK STANDARD; PRT; 528 AA.
AC P09480;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=88283624; PubMed=3267226;
RA Nef P., Oneysen C., Alliod C., Couturier S., Ballivet M.;
RT "Genes expressed in the brain define three distinct neuronal
RL nicotinic acetylcholine receptors."
RL EMBO J. 7:595-601(1988).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different
CC type of subunits: alpha and non-alpha (also called beta). A
CC functional receptor seems to consist of two alpha-chains and
CC three non-alpha chains.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:30:03 ; Search time 48.3701 Seconds
(without alignments)
5022.709 Million cell updates/sec

Title: US-09-303-232-2
Perfect score: 4043
Sequence: 1 MKNAQLKTEVDDDELWLV.....MEALLATIAVLISAPHIIVS 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25: *
1: sp.archaea: *
2: sp.bacteria: *
3: sp.fungi: *
4: sp.human: *
5: sp.invertebrate: *
6: sp.mammal: *
7: sp.mhc: *
8: sp.organelle: *
9: sp.phage: *
10: sp.plant: *
11: sp.rodent: *
12: sp.virus: *
13: sp.vertebrate: *
14: sp.unclassified: *
15: sp.virus: *
16: sp.bacteriaph: *
17: sp.archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4014.5	99.3	807	5	Q8T7V5 drosophila
2	2613	64.6	570	5	Q9VJT9 drosophila
3	2388.5	59.1	482	5	Q8T5F5 drosophila
4	1991	49.2	542	5	Q86NN7 drosophila
5	1988	49.2	545	5	Q9VW19 drosophila
6	1850	45.8	496	5	Q9XZ13 drosophila
7	1625.5	40.2	494	5	Q8T7S2 drosophila
8	1622.5	40.1	494	5	Q8T7S3 drosophila
9	1621.5	40.1	494	5	Q86MN8 drosophila
10	1620.5	40.1	494	5	Q8T7S1 drosophila
11	1611	39.8	523	5	Q8T7R9 drosophila
12	1609	39.8	501	5	Q9XZ14 drosophila
13	1602	39.6	509	5	Q8T7S0 drosophila
14	1502	37.2	391	5	Q9NKG2 drosophila
15	1310.5	32.4	554	5	Q9V179 drosophila
16	1235	30.5	525	5	Q8IPE2 drosophila

RESULT 1

ID	Q8T7V5	PRELIMINARY;	PRT;	807 AA.
AC	Q8T7V5;			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DE	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Nicotinic acetylcholine receptor Dalpha5 subunit			
GN	NACR-ALPHA-34E OR NACRALPHA-34E OR BG:DS05899.4 OR BG:DS05899.5 OR CG4498 OR CG16878 OR CG32975.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21969411; PubMed=11973307;			
RA	Grauso M., Reenan R.A., Culotto E., Sattelle D.B.;			
RT	"Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing.";			
RL	Genetics 160:1519-1533(2002).			
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.			
DR	EMBL; AF272778; AM13390.1; -			
DR	FlyBase; FBgn0028875; NACR-alpha-34E.			
DR	GO; GO:0016021; C:intracel to membrane; IEA.			
DR	GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.			
DR	GO; GO:0005216; F:ion channel activity; IEA.			
DR	GO; GO:0030594; F:neurotransmitter receptor activity; IEA.			
DR	GO; GO:0006811; P:ion transport; IEA.			
DR	GO; GO:0007268; P:synaptic transmission; IEA.			
DR	InterPro; IPR006201; Neur_chan.			
DR	InterPro; IPR006202; Neur_chan_LBD.			
DR	InterPro; IPR006029; Neur_chan_memb.			
DR	Pfam; PF029311; Neur_chan_LBD; 1.			
DR	Pfam; PF02932; Neur_chan_memb; 1.			

ALIGNMENTS

17	1197	29.6	509	13	Q800C7
18	1185.5	29.3	502	11	Q9JHD6
19	1174	29.0	513	13	Q7T2R9
20	1168	28.9	511	13	Q03481
21	1163	28.8	555	13	Q7T2U0
22	1156.5	28.6	502	6	Q866A2
23	1147.5	28.4	486	13	Q7T2S0
24	1140	28.2	335	5	Q9NKG1
25	1108.5	27.4	554	13	Q7T2T9
26	1072.5	26.5	474	13	Q7T2U1
27	1015	25.1	480	5	Q8I932
28	1009.5	25.0	461	5	Q9I197
29	977	24.2	554	5	Q62083
30	968.5	24.0	542	5	Q18556
31	902	22.3	537	5	Q8MUR0
32	897	22.2	515	5	Q46133
33	888.5	22.0	537	5	Q9U941
34	883.5	21.9	523	5	Q46128
35	882.5	21.8	552	5	Q9I765
36	876	21.7	568	5	Q9NFR5
37	875	21.6	545	5	Q96631
38	861.5	21.3	531	5	Q96632
39	859	21.2	536	5	Q8T0Y9
40	854	21.1	540	5	Q46134
41	854	21.1	595	5	Q9I764
42	851	21.0	509	5	Q9NFX8
43	849	21.0	497	5	Q46135
44	849	21.0	509	5	Q8MM21
45	849	21.0	529	13	Q7T2S4

Q800C7 brachydanio
Q9JHD6 mus musculus
Q7T2R9 fugu rubrip
Q03481 gallus gall
Q7T2U0 fugu rubrip
Q866A2 macaca mula
Q7T2S0 fugu rubrip
Q9NKG1 drosophila
Q7T2T9 fugu rubrip
Q7T2U1 fugu rubrip
Q8I932 caenorhabdi
Q9I197 caenorhabdi
Q62083 caenorhabdi
Q18556 caenorhabdi
Q8MUR0 apis mellif
Q46133 locusta mig
Q9U941 myzus persi
Q46128 heliothis v
Q9I765 myzus persi
Q9NFR5 drosophila
Q96631 heliothis v
Q96632 heliothis v
Q8T0Y9 alysia cal
Q46134 locusta mig
Q9I764 myzus persi
Q9NFX8 myzus persi
Q46135 locusta mig
Q8MM21 aphid gossy
Q7T2S4 fugu rubrip

DR PRINTS: PR00252; NRIONCHANNEL.
DR TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 807 AA; 91223 MW; C8B4F6B34287C8C8 CRC64;

Query Match 99.3%; Score 4014.5; DB 5; Length 807;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 770; Conservative 0; Mismatches 0; Indels 37; Gaps 1;

QY 1 MKNAQLKLTVEVDDDELAVLRLAHCSSNFSSSSSTRITSSNQNRHQQLTLQPRSLSTKH 60
Db 1 MKNAQLKLTVEVDDDELAVLRLAHCSSNFSSSSSTRITSSNQNRHQQLTLQPRSLSTKH 60
QY 61 HSNIASQHNSSQOQEPASKDEVDVANHGSRNDQOHLQOLDSSNMLSPKTAATAAGDEA 120
Db 61 HSNIASQHNSSQOQEPASKDEVDVANHGSRNDQOHLQOLDSSNMLSPKTAATAAGDEA 120
QY 121 TQOQPTNIRLCARQRLRRRRKKPKATNETDIKKOQLSMPPFKTRKSTDTYSTPAAT 180
Db 121 TQOQPTNIRLCARQRLRRRRKKPKATNETDIKKOQLSMPPFKTRKSTDTYSTPAAT 180
QY 181 TSQPTATYMOCRASDNFSPISRHDRVSTATFAWLHLVQLVLSLQOQLHVQQRSVL 240
Db 181 TSQPTATYMOCRASDNFSPISRHDRVSTATFAWLHLVQLVLSLQOQLHVQQRSVL 240
QY 241 LFRRIAATAFTSYLGSAQAQKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300
Db 241 LFRRIAATAFTSYLGSAQAQKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300
QY 301 LSAKVCAGVHEKRLHDLDPYNTLPRVNLNEDPLQLSGFLTMQIIVDEKNQLVLT 360
Db 301 LSAKVCAGVHEKRLHDLDPYNTLPRVNLNEDPLQLSGFLTMQIIVDEKNQLVLT 360
QY 361 NVWLKLEWMDNLRWNTSDYGGVKDLRIPPHRIWKPDLVLYNSADEGFGDTGYQTNVVVRN 420
Db 361 NVWLKLEWMDNLRWNTSDYGGVKDLRIPPHRIWKPDLVLYNSADEGFGDTGYQTNVVVRN 420
QY 421 NGSCLYVPPGIFKSTCKIDITWPFDDQRCMEKFGSWTYDGFQDLQLODETGGDISYV 480
Db 421 NGSCLYVPPGIFKSTCKIDITWPFDDQRCMEKFGSWTYDGFQDLQLODETGGDISYV 480
QY 481 LNSEWELLGVGKNEIYVNCPPPYDITFAIIRRTLYYRPNLIIPCVLIASMLLG 540
Db 481 LNSEWELLGVGKNEIYVNCPPPYDITFAIIRRTLYYRPNLIIPCVLIASMLLG 540
QY 541 FTLPDPSGKLSLGVTLISLTPLNVAETMPATSDAVPL----- 581
Db 541 FTLPDPSGKLSLGVTLISLTPLNVAETMPATSDAVPLTYFNCIMFMVASSVVST 600
QY 582 -----WIRIVELWLPILRMRGRPLILEFPTPCSDTSSERKHQ 623
Db 601 ILINLYHNRADTHEMSEWIRIVELWLPILRMRGRPLILEFPTPCSDTSSERKHQ 660
QY 624 ILSDVELKRSKSLANVLIDDDFRHNCRPMTPGGTLPHPAFYTYVGGQDGGSIGP 683
Db 661 ILSDVELKRSKSLANVLIDDDFRHNCRPMTPGGTLPHPAFYTYVGGQDGGSIGP 720
QY 684 IGSTRMPDVAVHTFCIKSSTEYELGLIKRIFITDQLRKDDCEANDKWKFAAMVVD 743
Db 721 IGSTRMPDVAVHTFCIKSSTEYELGLIKRIFITDQLRKDDCEANDKWKFAAMVVD 780
QY 744 LCLIFTWFAIATIAVLISAPHLIVS 770
Db 781 LCLIFTWFAIATIAVLISAPHLIVS 807

RESULT 2
Q9VJT9 PRELIMINARY; PRT; 570 AA.
AC Q9VJT9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE C32975-PA.
GN NACRALPHA-34B OR CG32975.
OS Drosophila melanogaster [fruit fly].
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Mortman J.R., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houstoun K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodgett W., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;
RA "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003642; AAF53374.3; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.

```
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 2.
DR Pfam; PF02932; Neur_chan_LBD; 2.
DR PRINTS; PS00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
SQ SEQUENCE 570 AA; 64767 MW; DBE2E1E185263BF CRC64;

Query Match 64.6%; Score 2613; DB 5; Length 570;
Best Local Similarity 69.0%; Pred. No. 1e-219;
Matches 532; Conservative 17; Mismatches 20; Indels 202; Gaps 9;

QY 1 MKNAQLKTEVDDDELWLAFLAHCSNFSSSSSTRITSSNQRINQQLTTIQPRSLSTKH 60
Db 1 MKNAQLKTEVDDDELWLAFLAHCSNFSSSSSTRITSSNQRINQQLTTIQPRSLSTKH 60

QY 61 HSNIASEQHSNQOQEPASKDEVDVANGHRSNDQOQTHLQQLDSSNMLSPKTAATAAGDEA 120
Db 61 HSNIASEQHSNQOQEPASKDEVDVANGHRSNDQOQTHLQQLDSSNMLSPKTAATAAGDEA 120

QY 121 TQOQTNIRLCARKQRLRRRRKKPKATNETDIKKQQQLSMPPFKTRKSTDYSTPAAT 180
Db 121 TQOQTNIRLCARKQRLRRRRKKPKATNETDIKKQQQLSMPPFKTRKSTDYSTPAAT 180

QY 181 TSCTPTATYMCRASDNESFIPISRHDRVSTATFAWLVHLVQLLVLSLQWQLHVQORSVL 240
Db 181 TSCTPTATYMCRASDNESFIPISRHDRVSTATFAWLVHLVQLLVLSLQWQLHVQORSVL 240

QY 241 LFRRIAASITAFISYLGSAFAQLKNSSSSSSS-SNSNNSNSTOILNGLNKHSMIFLLIYL 299
Db 241 LFRRIAASITAFISYLGSAFAQLKNSSSSSSS-SNSNNSNSTOILNGLNKHSMIFLLIYL 299

QY 300 NLSAKVCLAGYHEKRLHLLDLPYNTLERPVNLESDFLQSLFGLTLMQIIDVDEKNQLLV 359
Db 300 NLSAKVCLAGYHEKRLHLLDLPYNTLERPVNLESDFLQSLFGLTLMQIIDVDEKNQLLV 359

QY 360 TLNVWLEWDMNLRWNTSDYGVKDLRIIPPHRIWKPDVLMYNSADGFGDTGTQNVVVR 419
Db 360 TLNVWLEWDMNLRWNTSDYGVKDLRIIPPHRIWKPDVLMYNSADGFGDTGTQNVVVR 419

QY 372 NNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGFQV-----416
Db 372 NNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGFQV-----416

QY 480 VLNGEWELLVGPKRNEIYNNCCPEPYDITFAIIIRRTLYFFNLIIPCVLIIASMAIL 539
Db 480 VLNGEWELLVGPKRNEIYNNCCPEPYDITFAIIIRRTLYFFNLIIPCVLIIASMAIL 539

QY 540 GFTLPDPSGKLSGLVITLLSLTVFLNMVAETWPATSDAVPLIRIVFLCWLPMRMSR 599
Db 540 GFTLPDPSGKLSGLVITLLSLTVFLNMVAETWPATSDAVPLIRIVFLCWLPMRMSR 599

QY 470 GFTLPDPSGKLSGLT-----YFNCIMPVASSVVST-----502
Db 470 GFTLPDPSGKLSGLT-----YFNCIMPVASSVVST-----502

QY 600 PGRPLILEPTTPCSDTSERKHQILSDVELKERSKSLANVLVDIDDDFRHNCRPMTPG 659
Db 600 PGRPLILEPTTPCSDTSERKHQILSDVELKERSKSLANVLVDIDDDFRHNCRPMTPG 659

QY 503 ----LILNY-----507
Db 503 ----LILNY-----507

QY 660 GTLPHNPAFYRTVYGDDSGSIGPTGSTRMPDAVTHHTCIKSSTEYELGLILKEIRFITD 719
Db 660 GTLPHNPAFYRTVYGDDSGSIGPTGSTRMPDAVTHHTCIKSSTEYELGLILKEIRFITD 719

QY 508 -----HH-----RNADTHEM-----SE 519
Db 508 -----HH-----RNADTHEM-----SE 519

QY 720 QLRKDECDNDIANDWKFAAMVVDRLCLIIFTMPAIIATIAVLLSAPHIIVS 770
Db 720 QLRKDECDNDIANDWKFAAMVVDRLCLIIFTMPAIIATIAVLLSAPHIIVS 770

QY 520 WLRKDECDNDIANDWKFAAMVVDRLCLIIFTMPAIIATIAVLLSAPHIIVS 570
Db 520 WLRKDECDNDIANDWKFAAMVVDRLCLIIFTMPAIIATIAVLLSAPHIIVS 570

RESULT 3
Q8T5F5
ID Q8T5F5 PRELIMINARY; PRT; 482 AA.
AC Q8T5F5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
```

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalp4 subunit.
GN NACR-ALPHA-34E OR NACRALPHA34E OR BG:DS05899.4 OR BG:DS05899.5 OR
GN CG4498 OR CG16878 OR CG32975.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=21969411; PubMed=11973307;
GX Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RA "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalp4s, Dalp4a6 and Dalp4a7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AY036613; AAK67256.1; -.
DR FlyBase; FBgn028875; nacr-alpha-34E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007268; P:receptor transmission; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PS00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 482 AA; 54445 MW; A4B95B921CA33F05 CRC64;
```

```
Query Match 59.1%; Score 2388.5; DB 5; Length 482;
Best Local Similarity 98.3%; Pred. No. 3.6e-200;
Matches 457; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 1 MKNAQLKTEVDDDELWLAFLAHCSNFSSSSSTRITSSNQRINQQLTTIQPRSLSTKH 60
Db 1 MKNAQLKTEVDDDELWLAFLAHCSNFSSSSSTRITSSNQRINQQLTTIQPRSLSTKH 60

QY 61 HSNIASEQHSNQOQEPASKDEVDVANGHRSNDQOQTHLQQLDSSNMLSPKTAATAAGDEA 120
Db 61 HSNIASEQHSNQOQEPASKDEVDVANGHRSNDQOQTHLQQLDSSNMLSPKTAATAAGDEA 120

QY 121 TQOQTNIRLCARKQRLRRRRKKPKATNETDIKKQQQLSMPPFKTRKSTDYSTPAAT 180
Db 121 TQOQTNIRLCARKQRLRRRRKKPKATNETDIKKQQQLSMPPFKTRKSTDYSTPAAT 180

QY 181 TSCTPTATYMCRASDNESFIPISRHDRVSTATFAWLVHLVQLLVLSLQWQLHVQORSVL 240
Db 181 TSCTPTATYMCRASDNESFIPISRHDRVSTATFAWLVHLVQLLVLSLQWQLHVQORSVL 240

QY 241 LFRRIAASITAFISYLGSAFAQLKNSSSSSSS-SNSNNSNSTOILNGLNKHSMIFLLIYL 299
Db 241 LFRRIAASITAFISYLGSAFAQLKNSSSSSSS-SNSNNSNSTOILNGLNKHSMIFLLIYL 299

QY 300 NLSAKVCLAGYHEKRLHLLDLPYNTLERPVNLESDFLQSLFGLTLMQIIDVDEKNQLLV 359
Db 300 NLSAKVCLAGYHEKRLHLLDLPYNTLERPVNLESDFLQSLFGLTLMQIIDVDEKNQLLV 359

QY 360 TLNVWLEWDMNLRWNTSDYGVKDLRIIPPHRIWKPDVLMYNSADGFGDTGTQNVVVR 419
Db 360 TLNVWLEWDMNLRWNTSDYGVKDLRIIPPHRIWKPDVLMYNSADGFGDTGTQNVVVR 419

QY 420 NNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGFQV 464
Db 420 NNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGFQV 464
```

RESULT 4					
Q86MN7	ID	Q86MN7	PRELIMINARY;	PRT;	542 AA.
AC	DT	01-JUN-2003	(TrEMBLrel. 24, Created)		
DT	DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
DE	DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	DT	Nicotinic acetylcholine receptor subunit alpha7 precursor.			
GN		NACRALPHA-18C.			
OS		Drosophila melanogaster (Fruit fly).			
OC		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC		Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC		Ephydroidea; Drosophilidae; Drosophila.			
OX		NCBI_TaxID=7227;			
RN		[1]			
RP		SEQUENCE FROM N.A.			
RA		Millar N.S.;			
RL		Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.			
DR		EMBL; AJ545210; CAD86936.1; -.			
DR		GO; GO:0016020; C:membrane; IEA.			
DR		GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . . ; IEA.			
DR		GO; GO:0030594; F:neurotransmitter receptor activity; IEA.			
DR		GO; GO:0006811; P:ion transport; IEA.			
DR		InterPro; IPR006201; Neur_channel.			
DR		InterPro; IPR006202; Neur_chan_LBD.			
DR		InterPro; IPR006029; Neu_channel_memb.			
DR		Pfam; PF02931; Neur_chan_LBD; 1.			
DR		Pfam; PF02932; Neur_chan_memb; 1.			
DR		PRINTS; PR00252; NRIONCHANNEL.			
DR		TIGRFAMS; TIGR00860; LIC; 1.			
DR		PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.			
KW		Signal; Receptor.			
FT		SIGNAL	1	POTENTIAL.	
FT		CHAIN	36	542	
FT				NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT DALPHA7.	
SQ		SEQUENCE	542 AA;	60988 MW;	69D2C39746BB74D7 CRC64;
		Query Match	49.24%;	Score 1991;	DB 5; Length 542;
		Best Local Similarity	72.84%;	Pred. No. 2.7e-165;	
		Matches 382;	Conservative 43;	Mismatches 46;	Indels 54; Gaps 0
QY		295	LLIY-- --INFSAKVCLAGYHEKRLHLLDDPYNTLERPVLNESDPQLSPGLTLMQIIDV	351	
Db		19	MLVYGLGLIMIACAAGPEKRLHALLDNYSLSERPVVNESDPLQLSFGLTLMQIIDV	78	
QY		352	DEKNQLLVTVWLKLEWMDMLRWNTSDYGVGKDLIPPHRIWKDPVLMYNSADSGFGDT	411	
Db		79	DERNQLITNTWLKLEWMDMLRNWSSEFGVGVDLRIPHLRKLPDVLMYNSADSGFGDT	138	
QY		412	YQTNVVVRNGSLCYVPPIGFSTCKIDITWFFPDDQRCMKFGSWTYDGFOLDLOLQE	471	
Db		139	YATNVVVRNGSLCYVPPIGFSTCKIDITWFFPDDQRCMKFGSWTYDGFOLDLOLQE	198	
QY		472	TGGDISSVLNGEWELLVGPGRNEIYNCCPEPIDITFAIIIRRTLYFPFNLIIPCVC	531	
Db		199	AGGDISSFITNGRWDLVGPGRNEIYNCCPEPIDITFAIIIRKRTLYFPFNLIIVPCV	258	
QY		532	LIASMALLGFTLPDPSGEKLSLGVTILLSLTVFLNNVAETMPATSDAVPL-----	581	
Db		259	LIASMALLGFTLPDPSGEKLSLGVTILLSLTVFLNNVAETMPATSDAVPLKGKYNFCIME	318	
QY		582	-----WIRIVLCWLPWILMRSPGRPLIIIEFTTPCS	614	
Db		319	MVASVVSTILVINHYHRNPDTHEMSERIVFYLWLFCLIRMQRPGQ-VGYECPPPPSS	377	
QY		615	DTSS---ERKHQILDVELKERSSKILLANVLIDDDEFRHNCRPMPTPGQTLPHNPAFYRT	671	
Db		378	SSSSASGEKKQOI-QNVELKERSSKILLANVLIDDDEF-CNHRCASATLPHQPYRYET	434	
QY		672	VYQGQDGDSIGFIGST-----RMPDAVTHHTFCIKSFEYELGLILKEIRFTDQLRKDD	725	


```
Db 197 DIGSFVTNGEWELIGVGRKNEIYVCCPEPYIDITEAVVIRKTLVYFFENLIVPCVLLIA 256
Qy 535 SMALLGFTLPDPSGKSLGVTLLSLTVFLNMVAETMPATSDAVPL----- 581
Db 257 SMALLGFTLPDPSGKSLGVTLLSLTVFLNMVAETMPATSDAVPLGTYFNCIMFWA 316
Qy 582 -----WIRIVLCWLPWILRMSRPG---RPLILEPPTPCS 614
Db 317 SSVVSTILLNYYHRRHADTHMSDMIRCVFLWLPWILRMSRPGSATTPPPARVPPP-- 374
Qy 615 DTSSERKHQLSDVELKERSKSLANVLDDDDFRHNCRPMTPGGTLPHPNPAFYRTVYG 674
Db 375 -----DLELRERSKSLANVLDDDDFRH-----PQAQPOCCRYR---- 412
Qy 675 QGDGSGTIGSTRMPDAVTHHTCKSTSEYELGLILKEIRFITDQARKDCNDIANDW 734
Db 413 GGEENGAG-----LAHSCF--GVDYELSLILKEIRVITDQMKDKDEDADISRDW 460
Qy 735 KFAAMVVDRLCLIIFTFAILATIAVLISAPHIIVS 770
Db 461 KFAAMVVDRLCLIIFTFAILATIAVLISAPHIIVS 496

RESULT 7
Q8T7S2
ID Q8T7S2 PRELIMINARY; PRT; 494 AA.
AC Q8T7S2;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphag6 subunit variant type II.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
[1]
RN NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas, Dalphag6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321446; AAM13393.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:00030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02933; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTRP ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56048 MW; 6BE711810ED7B8B CRC64;

Query Match 40.2%; Score 1625.5; DB 5; Length 494;
Best Local Similarity 61.4%; Pred. No. 2.3e-133;
Matches 320; Conservative 47; Mismatches 75; Indels 79; Gaps 7;
Qy 291 SWIFLLIYLINLSAKVCLAGYHEKRLHLHLLDLPYNTLERPVLSNEDPLQLSGLTLMQIID 350
```

```
Db 10 SLFVLLIPLAIKESC--QGPHEKELLNHLSTYNTLERPVANESLEPKFGLTLQIID 68
Qy 351 VDEKNQLLVNWLKLEWMDNLRWNTSDYGVKDLRIPPHRIWKPDVLYNSADEGPDG 410
Db 69 VDEKNQLLVNWLKLEWMDNLRWNETEYGVKDLRITPNKWKPDVLYNSADEGPDG 128
Qy 411 TYQTNVVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCQEMKFGSWTYDGFOLDLQIQD 470
Db 129 TYHTNIVVHGSGCLYVPPGIFKSTCKMDITWFFDDQRCQEMKFGSWTYDGNQLDLVNS 188
Qy 471 ETGDDISSYVLNGWELLGVPGKNEIYVCCPEPYIDITEAVVIRKTLVYFFENLIPC 530
Db 189 EDGDLSDFITNGEWWYLLAMPKKNITVYACCPPEYVDITFTIQRRTLYFFENLIVPC 248
Qy 531 VLTIASMLLGLTLPDPSGKSLGVTLLSLTVFLNMVAETMPATSDAVPL----- 581
Db 249 VLTISSMLLGLTLPDPSGKSLGVTLLSLTVFLNMVAESMPTSDAVPLIGTYFNCIM 308
Qy 582 -----WIRIVLCWLPWILRMSRPGTLPHPNPAFY 613
Db 309 FVASSVVLTVVLNYYHRTADITHMPPWIKSVFLQWLPWILRMRGPRKI----- 359
Qy 614 SDTSSEKHKQLS-----DVELKERSKSLANVLDDDDFRHNCRPMTPGGTLPHPNPAFY 669
Db 360 -----TRKTIILSNRMEKELKERSKSLANVLDDDDFRHTI----- 398
Qy 670 RTVYVQGDGSGTIGSTRMPDAV--THHTCKSTSEYELGLILKEIRFITDQARKDCBN 728
Db 399 -----SGSQTALGSSASGRTTVEEHTAI--GCNKHDLHLILKELQFITARMKXADDEA 452
Qy 729 DIANDWKFAAMVVDRLCLIIFTFAILATIAVLISAPHIIV 769
Db 453 ELIGDWKFAAMVVDRLCLIVFTLTIATVTLISAPHIIV 493

RESULT 8
Q8T7S3
ID Q8T7S3 PRELIMINARY; PRT; 494 AA.
AC Q8T7S3;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphag6 subunit variant type I.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
[1]
RN NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas, Dalphag6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321445; AAM13392.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:00030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02933; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTRP ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56048 MW; 6BE711810ED7B8B CRC64;
```



```
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56095 MW; B46BEDA63A92942 CRC64;

Query Match 40.1%; Score 1622.5; DB 5; Length 494;
Best Local Similarity 61.2%; Pred. No. 4.1e-133;
Matches 319; Conservative 46; Mismatches 77; Indels 79; Gaps 7;

QY 291 SWIFLLIYLNLAKVCLAGVHEKRLHDLDPNTLERPVNLNEDPLQLSFGTLMOIID 350
Db 10 SLFVLLIFLAIKESC-QGPKERLLNHLSTNTLERPVANSEPLEVFGUTLQIID 68
QY 351 VDEKNQLLVNWLKLEWDMNLRWNTSDYGVKDLRIPPHRIWKPDVLMYNSADEGFDG 410
Db 69 VDEKNQLITNWLLEWDMNLRWNETEYGVKDLRITENKWKPDVLMYNSADEGFDG 128
QY 411 TYQTNVVRNNGSLVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFOLDIQLQD 470
Db 129 TYHTNIWVHNGSLVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGNQLDLVLS 188
QY 471 ETGGDISVYLVNGEWELLVGPKRNEYNCPEYDITFAIIIRRTLYYFFNLIIPC 530
Db 189 EDGDLSDFTNGEWYLLAMPKKNIVYACCEPVVDITFTIQRRTLYYFFNLIIPC 248
QY 531 VLIASWALLGFTLPDPSGKLSLGVTLISLTVFLNVAETMPATSDAVEL----- 581
Db 249 VLISSWALLGFTLPDPSGKLSLGVTLISLTVFLNVAETLPQVSDAIFLLGTYENCIM 308
QY 582 -----WIRIVFLCWLPIILMSRGRPLILEFFTPPC 613
Db 309 FMVASSVVLTVVLYNHYHRTADIHMPPIKSVFLQWLPILRMGRPKRI----- 359
QY 614 SDTSSSRKHQILS-----DVELKERSKSLLANVLIDDDDFRHNCRPMTGGTLPHNPAFY 669
Db 360 -----TRKYLISNRKKELEKERSKSLLANVLIDDDDFRHTI----- 398
QY 670 RTVYGQDGSIGPIGSTRMPDAV-THHTCIKSTSEYELGLILKEIRFITDQLRKDECN 728
Db 399 -----SGSQTAIGSSASFGRPITVEEHHTAI-GCNHKDLHLILKEIQFITARMKADDEA 452
QY 729 DIANDKFAAMVVDRLCLIFTFMFAIATAVLLSAPHIIV 769
Db 453 ELIGDKWKAAMVVDRLCLIVFTLFTIATVTVLLSAPHIIV 493

RESULT 9
Q86MN8 PRELIMINARY; PRT; 494 AA.
AC Q86MN8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor subunit Dalpna6 precursor.
GN NACRALPHA-30D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Millar N.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ554209; CAD86935.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_channel.
```

RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing.";
RT Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321447; AAM13394.1; -.
DR FlyBase; FBgn0032151; nacr-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005216; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0007268; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRfams; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56113 MW; 48327537229573FF CRC64;

Query Match 40.1%; Score 1620.5; DB 5; Length 494;
Best Local Similarity 61.0%; Pred. No. 6.2e-133; Indels 79; Gaps 7;
Matches 310; Conservative 49; Mismatches 75;

QY 291 SWIFLLIYLNLAKVCLAGVHEKRLHLLDPYNTLERPVNLSDPLQLSGLTLMQIID 350
Db 10 SLFVLLIFLAIKESC-QGPEKRLNHLSTYNTLERPVANSEPLEVFGTLQIID 68

QY 351 VDEKNQLLVTVNWLKLEWDMNLWNTSDYGVKDLRIPIPHRIWKPDVLMYNSADEGFDG 410
Db 69 VDEKNQLLITNMLSLWMDNLYRNWNETYGGVKDLRIITPNKMLKPDVLMYNSADEGFDG 128

QY 411 TYOTNVVRNNGSLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFOLDLQD 470
Db 129 TYHTNVVVKHSGSLYVPPGIFKSTCKMDITWPFDDQHCMPKFGSWTYDGNQDLVNS 188

QY 471 ETGDISSYVNLGWEWELLGVPGKREIYVNCCEPYDITFAIIRRTLYFFNLIPIC 530
Db 189 EDGDLSDFITNGEYLLAMPKKNITVYACPEPYVDITFTQIRRTLYFFNLIVPC 248

QY 531 VLASALLGFTLPDPSGEKLSGLVTLLSLTVFLNVAETMPATSDAVPL----- 581
Db 249 VLISALLGFTLPDPSGEKLTGLVTLLSLTVFLNVAETMPATSDAVPLIGVTLLSL 308

QY 582 -----WIRIVFLCWLPLRLMSRRGRPLILEFPTTC 613
Db 309 FMVASSVVLTVVNLVYHRTADIHEMPWIKSVFLQWLPLRMGRGRKI----- 359

QY 614 SDTSSEKXHOILS-----DVELKERSKSLLANVLIDDDFRHNCRPMTPGOTLPHNPAFY 669
Db 360 -----TRKTILLSNRMKELKERSKSLLANVLIDDDFRHTI----- 398

QY 670 RTVYGQDDSGISGPIGSTRMPDAV-THHTCIKSTVEYELGLILKEIRFIQDLRKDECN 728
Db 399 -----SGSQAIGSSAAGFRPTTVEEHTAI-GCNHKDLHLILKELOFITARMKRKADEA 452

QY 729 DIANDMKFAAMVVDRLCLIIFTWFAILATIIVLSAPHITV 769
Db 453 ELIGDMKFAAMVVDRLCLIVFTLTIATTVLLSAPHIIV 493

RESULT 11
Q877R9 PRELIMINARY; PRT; 523 AA.
AC Q877R9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphas6 subunit variant type V.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Eohydroida; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing.";
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321449; AAM13396.1; -.
DR FlyBase; FBgn0032151; nacr-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005216; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006203; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRfams; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 523 AA; 59110 MW; 1C200AF74F87F841 CRC64;

Query Match 39.8%; Score 1611; DB 5; Length 523;
Best Local Similarity 58.2%; Pred. No. 4.6e-132;
Matches 320; Conservative 46; Mismatches 76; Indels 108; Gaps 7;

QY 291 SWIFLLIYLNLAKVCLAGVHEKRLHLLDPYNTLERPVNLSDPLQLSGLTLMQIID 350
Db 10 SLFVLLIFLAIKESC-QGPEKRLNHLSTYNTLERPVANSEPLEVFGTLQIID 68

QY 351 VDEKNQLLVTVNWLKLEWDMNLWNTSDYGVKDLRIPIPHRIWKPDVLMYNSADEGFDG 410
Db 69 VDEKNQLITNMLNLEWMDNLYRNWNETYGGVKDLRIITPNKMLKPDVLMYNSADEGFDG 128

QY 411 TYOTNVVRNNGSLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFOLDLQD 470
Db 129 TYHTNVVVKHSGSLYVPPGIFKSTCKMDITWPFDDQHCMPKFGSWTYDGNQDLVNS 188

QY 471 ETGDISSYVNLGWEWELLGVPGKREIYVNCCEPYDITFAIIRRTLYFFNLIPIC 530
Db 189 EDGDLSDFITNGEYLLAMPKKNITVYACPEPYVDITFTQIRRTLYFFNLIVPC 248

QY 531 VLASALLGFTLPDPSGEKLSGLVTLLSLTVFLNVAETMPATSDAVPL----- 581
Db 249 VLISALLGFTLPDPSGEKLTGLVTLLSLTVFLNVAETMPATSDAVPLIGVTLLSL 308

QY 582 -----WIRIVFLCWLPLRLMSRRGRPLILEFPTTC 613
Db 309 FMVASSVVLTVVNLVYHRTADIHEMPWIKSVFLQWLPLRMGRGRKI----- 359

QY 614 SDTSSEKXHOILS-----DVELKERSKSLLANVLIDDDFRHNCRPMTPGOTLPHNPAFY 669
Db 360 -----TRKTILLSNRMKELKERSKSLLANVLIDDDFRHTI----- 398

QY 670 RTVYGQDDSGISGPIGSTRMPDAV-THHTCIKSTVEYELGLILKEIRFIQDLRKDECN 728
Db 399 -----SGSQAIGSSAAGFRPTTVEEHTAI-GCNHKDLHLILKELOFITARMKRKADEA 452

QY 729 DIANDMKFAAMVVDRLCLIIFTWFAILATIIVLSAPHITV 769
Db 453 ELIGDMKFAAMVVDRLCLIVFTLTIATTVLLSAPHIIV 493

RESULT 11
Q877R9 PRELIMINARY; PRT; 523 AA.
AC Q877R9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

```
Db 415 NVLDIDDDFRHTI-----SGSQTALSSASGRPTTVEEHTAI 453
QY 700 KSSTYELGILILKEIRITDQLRKDDCNIDANKWFAAMVVDRLCLLIIFTMFAILATIA 759
Db 454 -GCHNKLHLILKEQLTITARMKADAEALIGDWKFAAMVVDRLCLIVFTLFIATVT 512
QY 760 VLLSAPHIIV 769
Db 513 VLLSAPHIIV 522

RESULT 12
Q9XZ14
ID Q9XZ14 PRELIMINARY; PRT; 501 AA.
AC Q9XZ14;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative nicotinic acetylcholine receptor alpha 7-2 subunit.
OS Heliothis virescens (Noctuid moth) (Owlet moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Heliothinae; Heliothis.
OX NCBI_TaxID=7102;
RN [1]
RP SEQUENCE FROM N.A.
RA Schultze T., Oellers N., Adamczewski M.;
RT "Putative alpha subunits of insect nicotinic acetylcholine receptors
RT more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than
RT to other insect nicotinic acetylcholine receptor alpha subunits.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF143847; AAD32698.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
DR Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 501 AA; 56704 MW; 43CB0DC3960C78AB CRC64;

Query Match 39.8%; Score 1609; DB 5; Length 501;
Best Local Similarity 60.8%; Pred. No. 6.4e-132;
Matches 319; Conservative 57; Mismatches 65; Indels 84; Gaps 10;

QY 295 LLYLNLAKVCLAGVHEKRLHLDLPYNTLRLPVLNESPQLSGLTLMQIIDVDEK 354
Db 10 LLALLPVSEQ---GPHEKRLNALLANYNTLRLPVLNESPQLSGLTLMQIIDVDEK 65
QY 355 NQLLVNTNVLKRWNDNLRWNTSDYGVKDLRIPHRWKPDVLMVNSADEGFDGTQYT 414
Db 66 NQLLVNTNVLKRWNDNLRWNTSDYGVKDLRIPHRWKPDVLMVNSADEGFDGTQYT 125
QY 415 NVVVRNNGSLYVPPGIFKSTKIDITWFFDDQRCCEMKFGSWTYDGFOLDLQDBTGG 474
Db 126 NVVVRSGSLYVPPGIFKSTKIDITWFFDDQRCCEMKFGSWTYDGNQLDLVLEAGG 185
QY 475 DISSVLNGEWELLGVPGKNEIYNNCCPEYIDITFAIIRRTLYFFNLIIPCVLIA 534
Db 186 DLSDFITNGEWELIGMGKKNITTYACCPPEYVDVTFIMIRRTLYFFNLIIPCVLIS 245
```

```
QY 535 SMALLGFTLPDPSGEKLSGLVTLLSLTFLNVAETMPATSDAVPL----- 581
Db 246 SMALLGFTLPDPSGEKLSGLVTLLSLTFLNVAETLPQVSDAIPLLGTGYFNCIMFWA 305
QY 582 -----WIRIVLCWLPWILRMSRPGRLILRPPTTPCSDTS 617
Db 306 SSVVLTVVLYNYHRTADITHMPQWIKSVFLWLPWILRMSRPGKKTIRK---TIMNTR 362
QY 618 SERKHQILSDVELKERSKSLANVLDDDDFRHNCRPMPTG-----GTLPHNPAFYRTV 672
Db 363 -----MRELKERSKSLANVLDDDDFRHG---PPPNSTASTGNLPGCGSIERTD 413
QY 673 YQGQ-----DDGSIGPIGSTMPDPAVTHHCISKSTSEYELGLILKEIRFITDQLRKD 724
Db 414 FRFSVRPSTMEDVG--GGLGS-----ELHLILRELQFIRMKKA 455
QY 725 DECNDIANDWKFAAMVVDRLCLLIIFTMFAILATIAVLLSAPHIIV 769
Db 456 DEAEALISDWKFAAMVVDRLCLFVFTLFTIATVAVLLSAPHIIV 500

RESULT 13
Q8T7S0
ID Q8T7S0 PRELIMINARY; PRT; 509 AA.
AC Q8T7S0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type IV.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321448; AAM13395.1; -.
DR FlyBase; FBgn0032151; nAChR-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
DR Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 509 AA; 57887 MW; BE8D8E019850C2BD CRC64;

Query Match 39.6%; Score 1602; DB 5; Length 509;
Best Local Similarity 59.3%; Pred. No. 2.7e-131;
Matches 318; Conservative 48; Mismatches 76; Indels 94; Gaps 8;

QY 291 SWIFLLIYLNLAKVCLAGVHEKRLHLDLPYNTLRLPVLNESPQLSGLTLMQIID 350
Db 10 SLFVLLIYLNLAKVCLAGVHEKRLHLDLPYNTLRLPVLNESPQLSGLTLMQIID 68
```

QY 351 V-----DEKNOLLVTNVWLKLEWMDMLRWNTSDYGGVKOLRIPPHRIWK 395
DB 69 VDEKNQILTTNAWNLDKQKLLITNLWLSLEWMDYLNRWNETYGGVKOLRITPNKLNK 128
QY 396 PDVLMYNSADEGFGTGYOTNVVVRNNGSCLYVPPGIPKSTCKIDITWPPDDQRCCKMFG 455
DB 129 PDVLMYNSADEGFGTGYOTNVVVRNNGSCLYVPPGIPKSTCKIDITWPPDDQRCCKMFG 188
QY 456 SWTVYDFGLDLODETGDISVVLNGEWELLGVPGKNEIYVNCCEPYIDITFALII 515
DB 189 SWTVYDFGLDLODETGDISVVLNGEWELLGVPGKNEIYVNCCEPYIDITFALII 248
QY 516 RRRLLYYFFNLIIPCVLIASMAILGFTLPDPSGKLSIGVTLLSLVFLNVAETMPAT 575
DB 249 RRRLLYYFFNLIIPCVLIASMAILGFTLPDPSGKLSIGVTLLSLVFLNVAETLPQV 308
QY 576 SDAVPL-----WIRVFLCWLPIWLRMS 598
DB 309 SDAIPLLGTGFNCIMFVASSVVLTVVVLNVHHRRTADITHPPWIKSVFLQWLPIWLRMG 368
QY 599 RRGRLPILIEFTTCSSTSRKHQILS---DVELKERSKSLANVLDDDDFRHNCR 654
DB 369 RRGRLPILIEFTTCSSTSRKHQILS---DVELKERSKSLANVLDDDDFRHTI- 413
QY 655 PMTPGGTLPHPAFRTYVYGGDGSIGPIGSTRMPDAV-THHTCIKSSTEYELGLILKE 713
DB 414 -----SGSQTALGSSASFCRPTTVEEHHTAI-GCNHKLHLILKE 452
QY 714 IRFITDOLRKDECNIDIANWKFAAMVVDRICLIIFTFMFAILATIAVLSAPHIIV 769
DB 453 LQFITARMRKADDEAEILGDKFAAMVVDRICLIIFTFMFAILATIAVLSAPHIIV 508

RESULT 14

Q9NKKD2 ID Q9NKKD2 PRELIMINARY; PRT; 391 AA.
AC Q9NKKD2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN NACR-ALPHA-34E OR BG:DS05899.4 OR BG:DS05899.5 OR CG4498 OR CG16878
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=9403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R.A., Harris N.L., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Celnikier S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219 (1999).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX Celnikier S.E., Agbayani A., Arcana T.T., Baxter E., Blazej R.G.,
RA Butenoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomontan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003408; AAF44854.1; -.
DR FlyBase; FBgn028875; nacr-alpha-34E.
KW Hypothetical protein.
SQ SEQUENCE 391 AA; 43974 MW; E2AB465CF275E8C0 CRC64;
Query Match 37.2%; Score 1502; DB 5; Length 391;
Best Local Similarity 94.4%; Pred. No. 1e-122;
Matches 301; Conservative 5; Mismatches 11; Indels 2; Gaps 2;
QY 1 MKNAQLKLTVDDELWLAFLAHCHSSNFSSSSSTRSSNORHNOQLTLOPESLSTKH 60
DB 1 MKNAQLKLTVDDELWLAFLAHCHSSNFSSSSSTRSSNORHNOQLTLOPESLSTKH 60
QY 61 HSNIASQHNSSQOEPAKDEVDVANGHRSNDQOHLQDLSSNNMSPKTAATAAGDEA 120
DB 61 HSNIASQHNSSQOEPAKDEVDVANGHRSNDQOHLQDLSSNNMSPKTAATAAGDEA 120
QY 121 TQOQTNIRLCARQRLRRRRKPKATPNETDIKKOQQLSMPFFKTRKSTDTYSTPAAT 180
DB 121 TQOQTNIRLCARQRLRRRRKPKATPNETDIKKOQQLSMPFFKTRKSTDTYSTPAAT 180
QY 181 TSCTATYMOCRASDNFSPISRHRDVRVSTATFAWLHLVLQVLVLSQQQLHVQOQSVL 240
DB 181 TSCTATYMOCRASDNFSPISRHRDVRVSTATFAWLHLVLQVLVLSQQQLHVQOQSVL 240
QY 241 LFRRIASTIAFISYLGSAQKNSSSSSSS-SSNSNNSSSTQILNGINKHSWIFLLIYL 299
DB 241 LFRRIASTIAFISYLGSAQKNSSSSSSSSSSSSSSSSTQILNGINKHSWIFLLIYL 300
QY 300 NLSAK-VCLAGYHEKRLH 317
DB 301 NLSAKGECHLNYLECCMQH 319

RESULT 15

Q9VL79 ID Q9VL79 PRELIMINARY; PRT; 554 AA.
AC Q9VL79;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CG4128 protein.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Shandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Perrier S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 11:50:34 ; Search time 5376.49 Seconds
(without alignments)
4276.744 Million cell updates/sec

Title: US-09-303-232-2
Perfect score: 4043
Sequence: 1 MKNAQLKLTVEVDDDELWLAV.....MFALLATIAVLVLSAPHIIVS 770

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cm2 1/USPTO.spool/US09303232/runat_07052004 101110 23893/app_query.fasta_1.2261
-DB=EST -QPM=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOP=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232 @CGN 1 1 4577 @runat_07052004 101110 23893 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	1231	30.4	885	12	BG632919	BG632919 GH16126.3
2	912	22.6	1201	9	AL530299	AL530299 AL530299
3	873.5	21.6	922	13	BU915857	BU915857 AGENCOURT
4	848.5	21.0	2296	14	CD013901	CD013901 90134548
5	837.5	20.7	1436	29	AY402873	AY402873 Homo sapi
6	826.5	20.4	1864	11	AK053497	AK053497 Mus muscu
7	826.5	20.4	2916	11	AK051730	AK051730 Mus muscu
8	826.5	20.4	3126	11	AK080415	AK080415 Mus muscu
9	822.5	20.3	1436	29	AY402875	AY402875 Homo sapi
10	803	19.9	1442	29	AY402876	AY402876 Homo sapi
11	802.5	19.8	1374	29	AY406230	AY406230 Homo sapi
12	798	19.7	1454	29	AY402878	AY402878 Mus muscu
13	797	19.7	1374	29	AY406232	AY406232 Mus muscu
14	797	19.7	4290	11	AK029177	AK029177 Mus muscu
15	791	19.6	1442	29	AY402877	AY402877 Pan trogl
16	787.5	19.5	2010	11	AK080475	AK080475 Mus muscu
17	787	19.5	2940	11	AK034228	AK034228 Mus muscu
18	787	19.5	3230	11	AK083157	AK083157 Mus muscu
19	780	19.3	1781	14	CD013889	CD013889 90138814
20	778.5	19.3	4037	11	AK049722	AK049722 Mus muscu
21	778.5	19.3	4046	11	AK051742	AK051742 Mus muscu
22	770	19.0	908	13	BU149265	BU149265 AGENCOURT
23	757	18.7	4589	11	AK030464	AK030464 Mus muscu
24	748.5	18.5	1374	29	AY406231	AY406231 Pan trogl
25	742.5	18.4	607	9	AI292581	AI292581 GH15518.5
26	729.5	18.0	1436	29	AY402874	AY402874 Pan trogl
27	724.5	17.9	1101	29	CNS0001F	AL058211 Drosophil
28	718.5	17.8	683	12	BM639954	BM639954 170006876
29	718.5	17.8	1007	29	CNS000HU	BM073676 Drosophil
30	711.5	17.6	1466	14	CD013891	CD013891 90139069
31	698.5	17.3	3483	11	AK081254	AK081254 Mus muscu
32	691.5	17.1	1034	13	EX403124	EX403124 BX403124
33	689	17.0	1751	14	CD013905	CD013905 90134608
34	689	17.0	2110	14	CD013902	CD013902 90134640
35	689	17.0	2110	14	CD013903	CD013903 90134855
36	685	16.9	902	29	AY407184	AY407184 Homo sapi
37	683	16.9	2513	11	AK033068	AK033068 Mus muscu
38	682.5	16.9	978	29	CNS006F9	AL064281 Drosophil
39	677	16.7	902	29	AY407186	AY407186 Mus muscu
40	673.5	16.7	2257	11	AK017571	AK017571 Mus muscu
41	663	16.4	895	13	BX735506	BX735506 BX735506
42	653.5	16.2	615	14	CB149460	CB149460 K-EST0205
43	646.5	16.0	833	14	CB245337	CB245337 UI-M-FYO-
44	643.5	15.9	791	14	CF739825	CF739825 UI-M-HDO-
45	643.5	15.9	2074	11	AK087554	AK087554 Mus muscu

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS
DEFINITION
BG632919 885 bp mRNA linear EST 23-APR-2001
GH16126.3prime GH Drosophila melanogaster head pot2 Drosophila
melanogaster cDNA clone GH16126.3 similar to CG128: FBan0004128
, ion channel, located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.
ACCESSION
VERSION BG632919.1 GI:13758409
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 885)
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
 Lewis, S., and Rubin, G.M.
 BDGP/HMI Drosophila EST Project
 Unpublished (2001)
 Other ESTs: GH16126.5prime
 Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 Based upon the presence of a XhoI site followed by a run of 14 or
 more T residues at the beginning of the sequence, this clone was
 polyadenylated. The resulting Poly-T sequence has been removed. hit
 genomic AE003511: arm: X [18792641,19136447]
 estimated-cyto:18A3-18C6: 04/10/2001
 Plate: GH.161 row: C column: 2
 High quality sequence stop: 784.

FEATURES

source
 1..885
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="GH16126"
 /sex="male and female"
 /dev_stage="adult"
 /lab_host="DH5 - alpha"
 /clone_lib="GH Drosophila melanogaster head pot2"
 /note="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 pot2. Plasmid cDNA library."

ORIGIN

Alignment Scores:
 Pred. No.: 2,97e-100 Length: 885
 Score: 1231.00 Matches: 221
 Percent Similarity: 98.33% Conservative: 14
 Best Local Similarity: 92.47% Mismatches: 4
 Query Match: 30.45% Indels: 0
 DB: 12 Gaps: 0
 US-09-303-232-2 (1-770) x BG632919 (1-885)

QY 343 LeuThrLeuMetGlnIleAlaAspValAspGluLysAsnGlnLeuLeuValThrAsnVal 362
 |||||
 |||||
 DB 885 CTACACTCATGCAGTATTCGATGTGCACGAAGAATCAACTGCTTATACGAATATT 826
 |||||
 |||||
 QY 363 TrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyrGlyGly 382
 |||||
 |||||
 DB 825 TGGCTCAATGGATGGATGAACATATGAATCTTCGATGGAATTCGAGTGGTGGTGT 766
 |||||
 |||||
 QY 383 ValLysAspLeuArgileProHisArgileTrpLysProAspValLeuMetTyrAsn 402
 |||||
 |||||
 DB 765 GTGCGGGATCTCGAATCCGACATCCGATCGCTATGAAACCGGATGCTACTGATGACAA 706
 |||||
 |||||
 QY 403 SerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgAsnAngly 422
 |||||
 |||||
 DB 705 AGTGCGGACGAGGGCTTCGATGGAACGATACGACCAAAATGTGGTTCGCAATATGGG 646
 |||||
 |||||
 QY 423 SerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIleThrTrp 442
 |||||
 |||||
 DB 645 AGCTGCTGTAGTACCGCAGGTATTTAGTCAAGTGAAGTCAAGTCAAGTCAAGTCAAGT 586
 |||||
 |||||
 QY 443 PheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAspGlyPhe 462
 |||||
 |||||
 DB 585 TTTCATTTCGACGATCAGAGATGTGAATGAAATTTGTTCTGTGGACCTACGATGGGTTT 526
 |||||
 |||||
 QY 463 GlnLeuAspLeuGlnLeuGlnAspGluThrGlyAspIleSerSerTyrValLeuAsn 482
 |||||
 |||||
 DB 525 CAGTTGGACCTGCAGTTGCAGGACGAAGCTGGTGGCGACATTCTTACGCTTTATACCAAT 466
 |||||
 |||||

QY 483 GlyGluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCys 502
 |||||
 |||||
 DB 465 GCGAATGGGACITGTTAGTGTGCCCGGTAAACGAAATGAAATCTACTATAATTTGCTGC 406
 |||||
 |||||
 QY 503 ProGluProTyrIleAspIleThrPheAlaIleIleIleArgArgArgThrLeuTyrTyr 522
 |||||
 |||||
 DB 405 CCAGAACCTTATATTGACATAACATTGCCATTGTAAGGCGCAAAACGTTGTACTAT 346
 |||||
 |||||
 QY 523 PhePheAsnLeuIleIleProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThr 542
 |||||
 |||||
 DB 345 TTTTTCATTCGATGTGCGGTGCTACTGATCGCTCCATGGCAGTGTAGGGTTTACA 286
 |||||
 |||||
 QY 543 LeuProProAspSerGlyLysLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThr 562
 |||||
 |||||
 DB 285 CTGCCACGACAGTTCGTGTAAGAGCTTTCGTTGGAGTTACAATCTTATTATCGCTTACA 226
 |||||
 |||||
 QY 563 ValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeu 581
 |||||
 |||||
 DB 225 GTCTTCTCTCAACATGGTGGCGGAAACAAATGCCGCGACCTCCGATGCGGTACCGCTG 169
 |||||
 |||||

RESULT 2

AL530299 1201 bp mRNA linear EST 23-MAY-2003
 LOCUS AL530299 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DD007YP05 5-PRIME, mRNA sequence.
 ACCESSION AL530299
 VERSION AL530299.2 GI:31068132
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Li W.B., Gruber C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 13, 2001 this sequence version replaced gi:12793792.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7646.r For more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD007CH03QP1&cluster=7646.r>. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Faraday Avenue Genoscope sequence ID : CS0DD007CH03QP1.

FEATURES

Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD007YP05"
 /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
 Pred. No.: 2,92e-71 Length: 1201
 Score: 912.00 Matches: 169
 Percent Similarity: 75.77% Conservative: 53
 Best Local Similarity: 57.68% Mismatches: 59
 Query Match: 22.56% Indels: 12
 DB: 9 Gaps: 2
 US-09-303-232-2 (1-770) x AL530299 (1-1201)


```

QY 308 AlaGlyTyrHisGluLysArgLeuLeuHisAspLeuLeuLeuAspProTyrAsnThrLeuGlu 327
Db 187 GCAGCCAGTCCAGAGAGAGCTTTACAGAGAGCTGGTCAAGAACTCAATCCCTTGGAG 246
QY 328 ArgProValLeuLeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGln 347
Db 247 AGGCCCGTGGCCCAATGACTCGCAACCACTACCCGTACTTCTCCCTGAGCCTCCTGCAG 306
QY 348 IleIleAspValLeuLeuAsnGlnLeuLeuValThrAsnValThrLeuLeuLeuGlu 367
Db 307 ATCATGACGCTGGATGAGAGAACCAAGTTTAAACCAACCAACATTTGGCTGCAATGTCT 366
QY 368 TrpAsnAspMetAsnLeuArgTyrAsnThrSerAspTyrGlyGlyValLysAspLeuArg 387
Db 367 TGGACAGATCACATTTTACAGTGAATGTCTCAGAAATCCAGGGGTGAGACTGTTCGT 426
QY 388 IleProHisArgIleThrLysProAspValLeuMetTyrAsnSerAlaAspGluGly 407
Db 427 TTCCAGATGCCAGATTTGGAACCAACAGACATTTCTCTATAACAGTGTCTGATGCGC 486
QY 408 PheAspGlyThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrVal 427
Db 487 TTGAGCCCATTCACATCACTACGTGTGTGGAATTCCTTGGGCATTCGCCAGTWCCTG 546
QY 428 ProProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAsp 447
Db 547 CTTCCAGGCATATTCAGAGTCTCTGCTACATGATGACGCTGGTTTCCCTTTGATGTG 606
QY 448 GlnArgCysGluMetLysPheGlySerThrTyrAspGlyPheGlnLeuLeuAspLeuGln 467
Db 607 CAGCACTGCAAACTGAAGTTGGGTCTCTGTCTTACGGAGCTGTCTTGGATCTGCAG 666
QY 468 LeuGlnAspGluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluThrGluLeu 487
Db 667 ATCCAGAG-----GCAGATATCAGTGGGTATATCCCAATCGAGATGGGACCTA 717
QY 488 LeuGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIle 507
Db 718 GTGGGAATCCCGCAGAGGAGTGAAGGTTCTATGAGTKTGTGCAAGAGCCCTACCCC 777
QY 508 AspIleThrPheAlaIleIleLeuArgArgThrLeuTyrTyrPhePheAsnLeuIle 527
Db 778 GATGTCACTTCAAGTACAGTACGCGCGCAGGAGCGCTCTACTATGGCCTCAACCTGCTG 837
QY 528 IleProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProAspSer 547
Db 838 ATCCCTGTGTCTATCTCCGCCCTCGCCCTCTGTGTCTCTGTCTCTCTGAGATTC 897
QY 548 -GlyGluLysLeuSerLeu-GlyValThrIleLeuLeuSerLeuThrValPheLeuAsnM 567
Db 898 GGGGAGAGATTTCCCTKGGGATACAGTYTTACTCTCTTACCGTCTTCATGCTGTC 957
QY 567 etValAlaGluThrMetProAlaThrSerAspAlaValProLeuThrIleArg----- 584
Db 958 TCGTGTCTGAGATCATGCGCGCAACATCGATTCGGTACCAITGAWAGCCAGTCTCTCG 1017
QY 585 -----IleValPheLeuCysTrpLeu 591
Db 1018 CCAGCACCATGATCATCGTGGGCTTTGGGGTGGTTG 1052

RESULT 3
BU915857
LOCUS
DEFINITION AGENCOURT 10492745 NICHD_XGC_001 Xenopus laevis cDNA clone
IMAGE:6642638 5', mRNA sequence.
ACCESSION BU915857
VERSION BU915857.1 GI:24097771
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.

```

```

REFERENCE 1 (bases 1 to 922)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgap@f-mail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM14228 row: f column: 14
High quality sequence stop: 746.
FEATURES
source
1..922
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6642638"
/tissue_type="oocytes"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_XGC_001"
/notes="vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.2 kb. Constructed by Life Technologies."
ORIGIN
Alignment Scores:
Pred. No.: 5,82e-68 Length: 922
Score: 873.50 Matches: 158
Percent Similarity: 75.09% Conservative: 50
Best Local Similarity: 57.04% Mismatches: 43
Query Match: 21.61% Indels: 26
DB: 13 Gaps: 4
US-09-303-232-2 (1-770) x BU915857 (1-922)

```

```

QY 352 AspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMet 371
Db 40 GATGAAAAAATCAGTACTTAACAACATATGCTCAAAATGCAATGATGATATAC 99
QY 372 AsnLeuArgTrpAsnThrSerAspTyrGlyValLysAspLeuArgIleProHis 391
Db 100 TATTTGCAGTGGAAACATGTCTGAATACCCCTGGGGTAAAAAATGTTCTTCAGATGGA 159
QY 392 ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db 160 CAGATTGGAAAGCCTGATATTCTCTCTATAACAGTGCAGATGATGATGATGCTCTCC 219
QY 412 TyrGlnThrAsnValValArgAsnGlySerCysLeuTyrValProProGlyIle 431
Db 220 TTTTACTATAATGACTGTGTGACTCCAGTGCAGCTGCAATATATCTCCATGATGAT 279
QY 432 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 451
Db 280 TTCAAAGTTCGTGTATACATAGATGCGATGGTTCCTTTTTCACATACAGAAATGTAAC 339
QY 452 MetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlu 471
Db 340 CTGAAGTTGGCTCTTGGACCTATGTCACCTGACCTGACCTGACCTGACCTGACCTG 396
QY 472 ThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuValPro 491
Db 397 -----TCAGACATATCTCAATATATCGTAATGAGAGTGGGACCTTGTAGCGTTCTC 450
QY 492 GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
Db 451 GGAAGAAAGAGTGAAGATTAATCTATCAATGCTCAAGGACCGTACCTGATGATGATGAT 510
QY 512 AlaIleIleIleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysVal 531

```

```

Db      511 ACAGTAAATAGAGCCAGAGACTCTATATACAGCACTTAACCTTCTAAATTCCTCTGTGTA 570
Qy      532 LeuIleAlaSerMetAlaLeuLeuLeuGlyPheThrLeuProAspSerGlyGluLysLeu 551
Db      571 CTGATATCTGCTCTAGCCCTGCTGGGTCTCTGCTCCAGCAGACTCAGGAGAGAAGATC 630
Qy      552 SerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThr 571
Db      631 TCACCTGGTATAACAGTCCTATTGTCCTCCTACCTGTCTGTTGAGTGGTAGTGAGATT 690
Qy      572 MetProAlaThrSerAspAlaValProLeu-----TTPIL 583
Db      691 ATGCCAGAACCTCAGACTCTGCTCCACTAATAGCAGACTTTGCCAGCACAATGGGT 750
Qy      583 e-----ArgIleValPheLeuCysTyrLeu----- 591
Db      751 TATTGTTGGGATTGTCAAGTGGTGTGTACAAAGTTGCTGGTCTTCAATATCATCAGCAT 810
Qy      592 -----ProTyrIleLeuArgMetSerArgProGlyArgProLeuIle 605
Db      811 GGATCCGCCCACTGGGTGGGAAATGCCCAAATGGGACCTAGATAAATT 859

```

RESULT 4

```

LOCUS      CD013901          2296 bp      mRNA      linear      EST 21-OCT-2003
DEFINITION 90134548 Single gene library Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD013901
VERSION    CD013901.1 GI:37777431
KEYWORDS   EST.
SOURCE     Homo sapiens (human)

```

ORGANISM

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Jin,P., Fu,G.K., Wilson,A.D., Yang,J., Chien,D., Hawkins,P.R.,
Au-Young,J. and Stuve,L.L.
PCR isolation and cloning of novel splice variant mRNAs from known
drug target genes
unpublished (2003)
Contact: Jin, P.
Incyte Corporation
3160 Porter Drive, Palo Alto, CA 94304, USA
Tel: 650 621 8639
Fax: 650 621 8965
Email: pjine@incyte.com.

```

JOURNAL

COMMENT

```

/clone_lib="Single gene library"
/clone_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Single gene library"
/clone_type="mRNA"
/db_xref="taxon:9606"

```

FEATURES

source

```

1..2296
    Location/Qualifiers
        organism="Homo sapiens"
        db_xref="taxon:9606"
        clone_lib="Single gene library"
        clone_type="mRNA"
        db_xref="taxon:9606"
        note="Vector: pDrive Cloning Vector; RT-PCR was performed
        using gene-specific primers flanking the open-reading
        frame. PCR products were subcloned into pDrive Cloning
        Vector and sequenced completely using M13 forward and
        reverse primers. Sequencing gaps were closed by
        re-sequencing using primers flanking the gapped areas."

```

ORIGIN

Alignment Scores:

```

Pred. No.: 4,25e-65      Length: 2296
Score:      848.50      Matches: 183
Percent Similarity: 54.11%      Conservative: 87
Best Local Similarity: 36.67%      Mismatches: 154
Query Match: 20.99%      Indels: 75
DB: 14      Gaps: 10

```

US-09-303-232-2 (1-770) x CD013901 (1-2296)

Qy 309 GlyTyrHis-----GluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThr 325

Db 332 GGCTCGCATACCGAGACTAGGAGCCGGCTTCAAAACACACCTCTTCGGGGGCTACACCGC 391

```

Qy      326 LeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeu 345
Db      392 TGGCGCGCCCGCGTCCCAACACTTCAGACGTGGTGATTGGCGCTTTGGACTCTCCATC 451
Qy      346 MetGlnIleIleAspValAspGluLysAsnGlnLeuLeuValThrAsnValTyrLeuLys 365
Db      452 GCTCAGCTCATCGATGTGGATGAGAGAACCAATATGATGACCAACCAACGCTGGCTAAAA 511
Qy      366 LeuGluTyrAsnAspMetAsnLeuLeuTyrAsnThrSerAspTyrGlyValLysAsp 385
Db      512 CAGGAGTGGAGCGCACTACAACTCGCTCGAAGCCCGCTGATTTTGCAACATCACATCT 571
Qy      386 LeuArgIleProProHisArgIleTyrLysProAspValLeuMetTyrAsnSerAlaAsp 405
Db      572 CTCAGGGTCCCTCTGAGATGATCGATCCCGACATTTGTTCTCTACACAATGCAGAT 631
Qy      406 GluGlyPheAspGlyThrTyrGlnThrAsnValValValArgAsnAsnGlySerCysLeu 425
Db      632 GGGGAGTTTGCAGTGACCCACATGACCAAGGCCACCTCTTCTCCACGGGCACTGTGCAC 691
Qy      426 TyrValProProGlyIlePheLysSerThrCysLysIleAspIleThrTyrPheProPhe 445
Db      692 TGGGTGCCCGCGCCATCTACAGAGCTCTCGACATCGACGTCACCTCTTCTCCCTTC 751
Qy      446 AspAspGlnArgCysGluMetLysPheGlySerTyrThrTyrAspGlyPheGlnLeuAsp 465
Db      752 GACCAGCAGAACTGCAGATGAAGTTTGGCTCTCTGGACTTATGACAGGCCCAAGATCGAC 811
Qy      466 LeuGlnLeuGlnAspGluThrGlyAspLysSerTyrValLeuAsnGlyGluTyr 485
Db      812 CTGAGCAGATGAGGAGACT---GTGGACCTGAGGAGACTACTGGGAGAGCGGCGAGTGG 868
Qy      486 GluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluPro 505
Db      869 GCCATCGTCAATGCCCGGCACCTACACAGCAGAGAGTACGACTGCTGCGCCGAGATC 928
Qy      506 TyrIleAspIleThrPheAlaIleIleIleArgArgArgThrLeuTyrTyrPhePheAsn 525
Db      929 TACCCCGACGTCACCTACGCTTCGTCATCGCGCGTGGCGCTCTTCTACACCATCAAC 988
Qy      526 LeuIleIleProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProPro 545
Db      989 CTCATCATCCCTGCTGCTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048
Qy      546 AspSerGlyGluLysLeuSerLeuGlyValThrIleLeuSerLeuThrValPheLeu 565
Db      1049 GACTGCGCGGAGAGATCACGCTGTGCATTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1108
Qy      566 AsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeu----- 581
Db      1109 CTGCTCATCTGAGATCATCCCGTCCACCTCGCTGGTGGTTCATCCGCTCATCGCGGAGTAC 1168
Qy      581 ----- 581
Db      1169 CTGCTGTTTACCATGATCTCTGCTCACCTGCTCATCGCTCATCGCTGCTGCTGCTGCTG 1228
Qy      582 -----TTPILeArgIleValPheLeu 588
Db      1229 GTGCACACACCGCTCCCGCAGCACCACACCATGCCCCATGGGTGGGCGGCGCTTCTG 1288
Qy      589 CysTyrLeuProTyrIleLeuArgMetSerArgProGlyArgProLeuLeuGluPhe 608
Db      1289 GGCTGTGTGCGCGCGTGGCTTCTGATGAACCGCGCGCCCGCCACCGCTGAGGCTC----- 1342
Qy      609 ProThrThrProCysSerAspThrSerSerGluArgLysHisGlnIleLeuSerAspVal 628
Db      1343 ---TGCCACCCCTTACGCTGAAAGTCAAGCCCTCTTATCATCTGCTGCTGCTGCTGCTGCTG 1399
Qy      629 GluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspLeuAspAsp 648
Db      1400 GATGCCGAGAGAGAGGAG-----GTGGTGTGGGAGGAGGAG 1435

```

```

QY 649 PheArgHisAsnCysArg---ProMetThrProGly---GlyThrLeuProHisAsnPro 666
Db 1436 GACAGATGGCATGTGCGAGGTGCGCCCTCTGCGGCACCCCTC----- 1483
QY 667 AlaPheTyrArgThrValTyrGlyGlnGlyAspGlySerIleGlyProIleGlySer 686
Db 1484 -----TGCAGCCACCGCCACCTCTGCGGCCTCAGGTCCCAAGGCTGAG 1531
QY 687 ThrArgMetProAspAla-----ValThrHisHisThrCysIleIysSerSerThr 703
Db 1532 GCTCTGCTGCGAGGGTGTGCTGCTATCACCCAC----- 1570
QY 704 GluTyrGluLeuGlyLeuIleLeuIysGluIleArgPheIleThrAspGlnLeuArgLys 723
Db 1571 -----ATGCAGAGCGCACTGAAGGTGTGCACTACATTCGCCACACCTGCGGTCT 1621
QY 724 AspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArg 743
Db 1622 GAGATGCTGACTCTTCGGTGAAGGAGGACTGGAAAGTATGTGCCATGTCATCGACAGG 1681
QY 744 LeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeu 762
Db 1682 ATCTTCCTCGCTGTTTATCATCTGCTGCTGCTGCGGACCATCGGCTCTTCTG 1738

RESULT 5
LOCUS AY402873 1436 bp DNA linear GSS 15-DEC-2003
DEFINITION Homo sapiens CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY402873
VERSION AY402873.1 GI:39759856
KEYWORDS GSS.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1436
/genes="CHRNA3"
/locus_tag="HCMI369"

gene
ORIGIN
Alignment Scores:
Pred. No.: 2,04e-64 Length: 1436
Score: 837.50 Matches: 172
Percent Similarity: 54.71% Conservative: 95
Best Local Similarity: 35.25% Mismatches: 158
Query Match: 20.71% Indels: 63
Db: 29 Gaps: 5

```

```

US-09-303-232-2 (1-770) x AY402873 (1-1436)
QY 312 GluLysArgLeuLeuHisAspLeuLeuAspP-oTyrAsnThrLeuGluArgProValLeu 331
Db 21 GAGCACCGCTCTATTGAGCGGCTGTTGAAGATTACATGATGATCATCGCGCTGTGGCC 80
QY 332 AsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuVal 351
Db 81 AACGTGTCTGACCCCATCATCATTCAGAGGTGTCCATGCTCAGCTGTGAAGGTG 140
QY 352 AspGluLysAsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMet 371
Db 141 GATGAAGTAAACCCAGATCATGAGACCAACTGTGGCTCAAGCAATCTGGAATGACTAC 200
QY 372 AsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProHis 391
Db 201 AAGCTGAAGTGAACCCCTCTGACTATGTGGGCGAGAGTTCATGCGTGTCCCTGCACAG 260
QY 392 ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db 261 AAGATCTGGAAGCCAGACATTTGCTGTATAACAATGCTGTGGGGATTTCAGGTGGAC 320
QY 412 TyrGlnThrAsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIle 431
Db 321 GACAGACCAAGCCTTACTCAAGTACACTGGGAGGTGACTTGGATACCTCCGCGCATC 380
QY 432 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 451
Db 381 TTTAAGAGCTCTCTGTAATAATCGAGTGCACCTACTTCCGTTTGATTACCAAACTGTACC 440
QY 452 MetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnAspGlu 471
Db 441 ATGAAGTTCGGTTCCTGCTCTACGATAAGCGGAAATCGATCTGCTGATCGGC--- 497
QY 472 ThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValPro 491
Db 498 TCCTTCCATGAACCTTCAAGGACTATTGGGAGAGCGCGGAGTGGGCGCATCATCAAGCCCCA 557
QY 492 GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
Db 558 GGCTACAAACACGACATCAAGTACACTGCTGCGAGAGATCTACCCCGCATCACATAC 617
QY 512 AlaIleIleLeuArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysVal 531
Db 618 TCCTGTACATCGCGCGCTGCGCTTGTCTACACCATCAACCTCATCATCCCCCTGCTG 677
QY 532 LeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeu 551
Db 678 CTCACTCTCCCTCACTGCTGCTGCTTCTTACTGCGCTCCCTCCGACTGCGGTGAGAGGTG 737
QY 552 SerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThr 571
Db 738 ACCCTGTGCATTTCTGTCTCTCTCTCCCTGACGCGTGTTCCTCTGCTGATCATGTAGACC 797
QY 572 MetProAlaThrSerAspAlaValProLeu----- 581
Db 798 ATCCCTTCCACCTCGCTGCTGCTATCCCGCTGAGAGATGACCTCTGCTGTTTCCACATGATT 857
QY 581 ----- 581
Db 858 TTTGTAACTTGTCCATGCTCATACCGCTTGTGTCCTCAACGTGCATACAGAACCCCG 917
QY 582 -----TrrIleArgIleValPheLeuCysTrpLeuProTrpIle 594
Db 918 ACGACACACACAATGCCCTCATGGGTGAGAGCTGATTCTTGAACCTGCTCCCGAGGTC 977
QY 595 LeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrThrProCysSer 614
Db 978 ATG-----TTCATGACGAGGCCAACAGC 1001
QY 615 AspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSer 634
Db 1002 AACGAGGCGACGCTCAGAAAGCGCGAGGCCCTCTACGGTGGCGGCTCTCAAAATCTGAAT 1061

```

QY	635	SerlyssSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCyArg	654
Db	1062	TGCTTCAGCCGCGAGAGTCC-----AAAGGCTGCAAG	1094
QY	655	ProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly	674
Db	1095	GAGGGTACCCCTCCAGGCGGATGTGGTTACTGCCACCCAGCAGGATAAAATC	1154
QY	675	GlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThr	694
Db	1155	TCCAATTTCAGTGTCACTACACGACGAGCTCTAGTCTGTAATCTGTTGATGCTGTGCTG	1214
QY	695	HisHisThrCysIleLeuSerSerThrGluTyrGluLeuGlyLeuIleLeuGlyGluIle	714
Db	1215	TCC-----CTCTGCTTGTGCACAGAAATCAAGAGCCATCAAGATGATTGG	1262
QY	715	ArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTyr	734
Db	1263	AAGTATATGCTGAATATGAAGCACAAATGAAGCAAGAGATTCAAGATGATTGG	1322
QY	735	LysPheAlaAlaMetValAspArgLeuCysLeuIleIlePheThrMetPheAlaIle	754
Db	1323	AAGTATGTTGCATGCTGATTGATGCTATTTTCTGTGGTTTTCACCTGCTGTGTCATT	1382
QY	755	LeuAlaThrIleAlaValLeuLeu 762	
Db	1383	CTAGGACAGCAGGATTGTTCTG 1406	
RESULT 6			
LOCUS	AK053497	1864 bp mRNA linear HTC 20-SEP-2003	
DEFINITION	Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30103E14 product:NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.		
ACCESSION	AK053497		
VERSION	AK053497.1	GI:26343494	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1. Carninci, P. and Hayashizaki, Y. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
QY	635	SerlyssSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCyArg	654
Db	1062	TGCTTCAGCCGCGAGAGTCC-----AAAGGCTGCAAG	1094
QY	655	ProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly	674
Db	1095	GAGGGTACCCCTCCAGGCGGATGTGGTTACTGCCACCCAGCAGGATAAAATC	1154
QY	675	GlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThr	694
Db	1155	TCCAATTTCAGTGTCACTACACGACGAGCTCTAGTCTGTAATCTGTTGATGCTGTGCTG	1214
QY	695	HisHisThrCysIleLeuSerSerThrGluTyrGluLeuGlyLeuIleLeuGlyGluIle	714
Db	1215	TCC-----CTCTGCTTGTGCACAGAAATCAAGAGCCATCAAGATGATTGG	1262
QY	715	ArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTyr	734
Db	1263	AAGTATATGCTGAATATGAAGCACAAATGAAGCAAGAGATTCAAGATGATTGG	1322
QY	735	LysPheAlaAlaMetValAspArgLeuCysLeuIleIlePheThrMetPheAlaIle	754
Db	1323	AAGTATGTTGCATGCTGATTGATGCTATTTTCTGTGGTTTTCACCTGCTGTGTCATT	1382
QY	755	LeuAlaThrIleAlaValLeuLeu 762	
Db	1383	CTAGGACAGCAGGATTGTTCTG 1406	
REFERENCE	Nature 409, 685-690 (2001)		
5	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
REFERENCE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
Direct Submission	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)		
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers		
FEATURES	source		
	1. 1864		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="FANTOM DB:EI30103E14"		
	/db_xref="MGI:2425288"		
	/db_xref="taxon:10090"		
	/clone="EI30103E14"		
	/tissue_type="eyeball"		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		
	/dev_stage="0 day neonate"		
	94. 1593		
	/note="unnamed protein product; NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT (SPTRIAL58471, evidence: FASTY, 99.8%ID, 100%length, match=1497) putative"		
	/codon_start=1		
	/protein_id="BAC35404.1"		
	/db_xref="GI:26343495"		
	/translation="MGVVLPPPLPMLMLVLLPVASASEHRLFQLFEDYNEIIL REVANSPHVIITQFEVMSQLVKVDENVQIMETNLKQIMNDYKXKSPSYQGVF MEVPAEKTKDPIDVLNNADGFDVDDTKKLLKTYTGPVTPPAIFKSSCKIDVTYF PDYQNCMTKFGWSYDKADIDLVIGSSMNLKYWESGEWAIKAPGYKHEIKYKNC BEIYQDIYSLYIRLPLPYTTLILPCLLISFLVLPFSPDCGKVTLCISVLS LTVFLVLTETIPSTLSVLPILCEYLLFTMIIVTISIVITVFLNVHVTPTTHMPT WKAVFLNLLPRVMTPTSTEEAPKTRNFYGAELSNLNCFSRADSKSKCEGIPQ DTCGYCHRRVKINFSANLRSRSSSDVLSLSPEIKASQVIAENNK AQWAKEIQDDWKYVAMVIDRFLWFLVILCVLTAGLFLQPLMARDDT"		
	1848. 1853		
	/note="putative"		
	1864		
	/note="putative"		
	polyA_signal		
	polyA_site		
	ORIGIN		
	Alignment Scores:		
	Pred. No.: 2,99e-63 Length: 1864		

Score:	826.50	Matches:	176
Percent Similarity:	54.04%	Conservative:	98
Best Local Similarity:	34.72%	Mismatches:	166
Query Match:	20.44%	Indels:	67
DB:	11	Gaps:	7

US-09-303-232-2 (1-770) x AK053497 (1-1864)

QY	295	LeuLeuIleTyrLeuAsnLeuSerAlaIysValCysLeuAlaGlyTyrHisGluLysArg	314
Db	127	ATGCTGATGCTGGTGTGATGCTGTCGCCAGTGGCCAGCGCTCGAAAGCTGAGCACCGC	186
QY	315	LeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSer	334
Db	187	CTGTTCCAGTACTCTGTTGAAGATTACACGAGATCATCCGGCCGGTGCTTAACGTGTCC	246
QY	335	AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleAspValaspGluLys	354
Db	247	CATCTGTGCATCATCCAGTTTGAGGTGTCATTGTCTCAGCTGGTGAAGGTGATGAAGTGA	306
QY	355	AsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArg	374
Db	307	AACCAAGATCATGGAAACCAACCTGTGGCTGAAGCAATCTGGAATGACTACAAGCTGAAA	366
QY	375	TrpAsnThrSerAspTyrGlyValIysAspLeuArgIleProProHisArgIleTrp	394
Db	367	TGGAAACCTCTGACTACCAAGGGGTGGAGTTTCATCGAGTCCCTCGACAGAAATCTGG	426
QY	395	LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr	414
Db	427	AAACCAAGATCTGGTCTTACCAACCCCGATGGGATTTCCAAAGTGGATGACAAACACC	486
QY	415	AsnValValAlaArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer	434
Db	487	AAAGCTCTACTCAAGTACACAGGAGAAGTGAAGTTCGGATCCCTCCGCCCATCTTTAAGAGC	546
QY	435	ThrCysIysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe	454
Db	547	TCATGCAAAATPCGATGTGACTACTTCCCGTTGACTACCAAAACTGCACCATGAAGTTC	606
QY	455	GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGly	474
Db	607	GGCTCTGTGCTACGACAAGCAAGATGACCTGTGCTCCTCATTTGGC---TCCTCCATG	663
QY	475	AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArg	494
Db	664	AACTCAAGGACTATTGGGAAAGTGGCGAGTGGCCATCATTAAGCCCGCGGTACAAA	723
QY	495	AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIle	514
Db	724	CATGAATCAAGTACAACTGCTGTGAGGAGATCTACCAAGACATCATCTACGTCTATAC	783
QY	515	IleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla	534
Db	784	ATTCCGCGCTGCCGTGTTCTACCATCAACTCATCTTCCTGTCCTCTCATCTTCC	843
QY	535	SerMetalLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly	554
Db	844	TTCCCTACTGTGCTCGTCTTCTACTCCCTCCGACTGTGGGAGAGGTGACGCTCTGC	903
QY	555	ValThrIleIleLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla	574
Db	904	ATTCCTGTCCTCTCCCTGACGGTCTTTCTCCTCGTGTGATCAACGAGACCATCCCTTC	963
QY	575	ThrSerAspAlaValProLeu-----	581
Db	964	ACCTCACTGGTCATCCCTTGATCGGGAGTAACTCCTCTTCACTATGATTTTGTCAAC	1023
QY	581	-----	581
Db	1024	TTGTCCATCGTCATCAGCTTTTGTGCTCAACGTGCACCTACAGAACTCCGACACACAC	1083
QY	582	-----TrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMet	597

QY 575 ThrSerAspAlaValProLeu----- 581
||||| :|||
Db 1040 ACCTCACGTGTCATCCCTTGATCGGGAGTACCTCTCTTCACTATGATTTTGTGTACC 1099
581
QY 1100 TTGTCCATCGTCATCACAGTCTTTGTGCTCAACGGTCACTACAGAACTCCGACCACACAC 1159
582 -----TtPlleArgIleValPheLeuCysTrpLeuProTtPlleLeuArgMet 597
|||:||||| :|||
Db 1160 ACGATGCCACTGGGTCAAGCTGTGTTTGAACCTTCTCCCGAGGTGATGTTTATG 1219
598 SerArgProGlyArgProLeuLeuGluPheProThrThrProCysSerAspThrSer 617
:||||| :|||
Db 1220 ACTAGGCCA-----ACGAGC 1234
618 SerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSer 637
:||||| :|||
Db 1235 ACCGAG-----GAACACGCCCAAGACGAGAACTTCTACGGTGCT 1276
638 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThr 657
|||:||||| :|||
Db 1277 GAGCTCTCAACACCTGAACTGCTTCAGCGTGCAGACTCCAAAAGCTGCAGGAAGGCTAC 1336
658 Pro-----GlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGlyGln 675
|||:||||| :|||
Db 1337 CCTGCCAAGATGGACATGTGGCTAT-----TGCACACCCGTCAGGTGTTAAATCTCA 1390
676 GlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThrHis 695
:||||| :|||
Db 1391 AATTTCAGTGCACACCTCACAGAAGCTCCAGTCTGTGATGCTGTGTGTGCTTGTCC 1450
696 HisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArg 715
|||:||||| :|||
Db 1451 -----CTGTCTGCTCTCTCCACCAGAAATCAAGAACGATCAAGATGTTGAAG 1498
716 PheIleThrAspGlnLeuArgLysAspGluCysAsnAspIleAlaAsnAspTrpLys 735
:||||| :|||
Db 1499 TATATTCTGCTGAAATATGAAGACACAGATGTACCCAAAGAGATCAAGATGTTGAAG 1558
736 PheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIleLeu 755
:||||| :|||
Db 1559 TATGTTGCCATGCTGATGATGATGATTTCTCTGGGTTTTCACCTCGGTGTCATTTA 1618
756 AlaThrIleAlaValLeuLeu 762
|||:||||| :|||
Db 1619 GGAACCTGCAGATTATTTCTG 1639

RESULT 8
AK080415
LOCUS
DEFINITION
Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:A730007P14 product:NEURONAL NICOTINIC
ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.
AK080415
VERSION
AK080415.1 GI:26348538
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3126)

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

TITLE
JOURNAL

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. .3126
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A730007P14"
/db_xref="MGI:2407635"
/db_xref="taxon:10090"
/clone="A730007P14"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
152. .1666
/note="unnamed protein product; NEURONAL NICOTINIC
ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT (STR)AAL58471,
evidence: FASTY, 99.8%ID, 100%length, match=1497)
putative"

FEATURES
source

CDS


```

Qy 734 TrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPheAla 753
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1255 TGGAGTACGTTGCAATGGTATGATGACCAACATCTCTCGAGTCTTCATGCTTGTGTGC 1314
Qy 754 IleLeuAlaThrIleAlaVal 760
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1315 ATCATCGGAACCTAGCGGTG 1335

RESULT 12
AY402878 1454 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus CHNB4 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402878
VERSION AY402878.1 GI:39759861
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1454)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1454)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1454
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1_>1454
/gene="CHNB4"
/locus_tag="HCM1370"

ORIGIN
Alignment Scores:
Pred. No.: 7,62e-61 Length: 1454
Score: 798.00 Matches: 175
Percent Similarity: 52.08% Conservative: 88
Best Local Similarity: 34.65% Mismatches: 152
Query Match: 19.74% Indels: 90
DB: 29 Gaps: 11

US-09-303-232-2 (1-770) x AY402878 (1-1454)

Qy 306 CysLeuAlaGlyTyrHisGluLysArgLeuLeuHisAspLeuLeuAspPro-----Tyr 323
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 6 TGCCGCTGGCCAGCAGAGGAGAGCTGTGGTATGCTCTCTGAACAAAACCTCGCTAC 65
Qy 324 AsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeu 343
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 66 AACAACTGATCCGCCAGCCACCGCTCTCCAGCTCATCTCCATCCGCTCGGCTGAGCTA 125
Qy 344 ThrLeuMetGlnIleIleAspValAspGluLysAsnGlnLeuValThrAsnValTrp 363
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 126 TCACTGTCCAGCTCATCATGTGATGAGCGAGACAGATCATGACCACAGCATCTCG 185
Qy 364 LeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyVal 383
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

```

```

Db 186 CTAATAACAGGAATGGACTGACTACCGCTGGCTGGAACAGCTCTCTGCTATGAAGGGTG 245
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 384 LysAspLeuArgIlePheProHisArgIleTrpLysProAspValLeuMetTyrAsnSer 403
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 246 AACATTCTGAGGATCTCTGCAAGCGAGTCTGGTCTGCTGACATGTGTTGTATCAACAAT 305
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 404 AlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValValArgAsnAsnGlySer 423
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 306 GCCGATGGGACCTATGAGGTGTCTGTCTACACCAACGATGATCGTGGTTCACACGGCAGC 365
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 424 CysLeuTyrValProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPhe 443
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 366 ATCCAGTGGTCCCTGCTCATCAAGAGTGCCTGCAAGATAGAGGTGAAGCACATT 425
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 444 ProPheAspAspGlnArgCysGluMetLysPheGlySerThrTyrAspGlyPheGln 463
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 426 CCCTTCGACCAAGCAACTGCACCTCAAAATTCGCTCTCGGACCTATGACCACACAGAG 485
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 464 LeuAspLeuGlnLeuAspGluThrGlyGlyAspIleSerSerTyrValLeuAsnGly 483
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 486 ATTGACATGGTCTTAATACACCCACA---GCCATCATGGATGACTTTACCCCGAGCGT 542
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 484 GluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysPro 503
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 543 GAATGGGACATGTGGCCCTCCGGGTGCAAGGACA-----GTGAACCCACAGGAC 593
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 504 GluProTyrIleAspIleThrPheAlaIleIleArgArgArgThrLeuTyrTyrPhe 523
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 594 CCCAGCTATGTGGACGTGACCTATCATCAAGCGACCACTCTTCTACACC 653
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 524 PheAsnLeuIleIleProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeu 543
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 654 ATCAATCTGATCATCCCTTGTGTCTCATCAGTCGCTGGCTATCTCTGCTCTTCTACCTG 713
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 544 ProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrVal 563
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 714 CCCTCCGACTGTGGGAGAGATGACGCTCTGCATCTCAGTCTGCTGGCTCATCATTC 773
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 564 PheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeu----- 581
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 774 TTCTGTGCTCATCTCCAGATCGTGGCTCCCACTCCCTTGCATTCCTCTCATTTGGC 833
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 581 ----- 581
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 834 AAGTACCTCTTGTTCACCATCGTGTGTCACTTTTCCATCGTCACCACTGTGTGTGTC 893
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 582 ----- 586
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 894 CTCATGTGCACACCGTTCCACCCAGCACTCACCATGGCATCTCTGGTCAAGGAGTGC 953
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 587 PheLeuCysTrpLeuProTrpIleLeuArgMetSerArgProGlyArgProLeuIleLeu 606
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 954 TTCTGCAACAGTGCCTCCACCTCTCTCTCATGAGCGCTCTGTT----- 998
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 607 GluPheProThrThrProCysSerAspThrSerSerGluArgLysHisGlnIleLeuSer 626
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 999 ---CTGGAAGTCAGCCCGCCGCGGTCCCTCATTCAGCCAGCTGCAC----- 1043
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 627 AspValGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAsp 646
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1044 ---TTGACCAACAGCTGGAAGTACGTCACCTCTGCTTGGTTCAGG----- 1082
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 647 AspAspPheArgHisAsnCysArgProMetThrProGlyGlyThrLeuProHisAsnPro 666
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1083 ----- 1118
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 667 AlaPheTyrArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySer 686
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1119 TACTTT-----GTGAACCTCTGCTCCCTGCC 1142
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Qy 687 ThrArgMetProAspAlaValThrHisThrCys-----Ile 699
    |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

```

```

Db      1143 ACT---CCTAGCTCTGCAGTCAGCTCCACACGGCAGGCGCTCCCGAGGATCCCGGGCTG 1199
Qy      700 LysSerSerThrGluTyr-----GluLeuGlyLeuLeuLeuGlyGluLeuArgPheLe 717
Db      1200 AGTCTCTCTGGAGGTTCCGCAAGATCTTACAGGAAGCATTAGAGGGGCTCAGCTTCATC 1259
Qy      718 ThrAspGlnLeuArgLysAspGluCysAsnAspIleAlaAsnAspTrpLysPheAla 737
Db      1260 GCACAGCATTGGAGAGTGATGATCGAGATCAGATGTCATCGAGAGCTGGAAATTCGTG 1319
Qy      738 AlaMetValValAspArgLeuCysLeuLeuLeuPheThrMetPheAlaIleLeuAlaThr 757
Db      1320 GCAATGGTCGTCACCGCTGTCTCTGTGGGTTCGTGATGTGTGTATCTGGGCACC 1379
Qy      758 IleAlaValLeuLeu 762
Db      1380 ATGGGGCTCTCTCTG 1394

RESULT 13
LOCUS   AY406232
DEFINITION Mus musculus CHRNAL gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY406232
VERSION   AY406232.1 GI:39762206
KEYWORDS  GSS.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1374)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE 2 (bases 1 to 1374)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE     Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source    1..1374
           /organism="Mus musculus"
           /mol_type="genomic DNA"
           /db_xref="taxon:10090"
           <1..>1374
gene      /gene="CHRNAL"
           /locus_tag="HCM2488"
ORIGIN
Alignment Scores:
Pred. No.:      8.6e-61      Length:      1374
Score:          797.00      Matches:    169
Percent Similarity: 51.97%      Conservative: 95
Best Local Similarity: 33.27%      Mismatches: 140
Query Matched: 19.71%      Indels:    104
DB:             29          Gaps:       8

US-09-303-232-2 (1-770) x AY406232 (1-1374)

Qy      291 SerTrpIlePheLeuLeuLeuTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr 310
Db      10 TCGACTGTCTCTGCTGCTAGGCTC---TGCTCCGCTGGCGCTTGTCTCTGGGCTCCGAA 66

```

```

Qy      311 HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal 330
Db      67 CATGAGACGCGTCTGGTGCAAGCTCTTTGAAGACTACAGCAGGTAGTCCGGCAGTG 126
Qy      331 LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAsp 350
Db      127 GAGGACCAACCGTGAGATTGTCAAGTCCCGTGGGTCTACAGCTGATCCAGCTTATCAAT 186
Qy      351 ValAspGluLysAsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp 370
Db      187 GTGGATGAAGTAATACAGATTGTGCACACCAATGTACCTCTGAAACAGCAATGGGTCCAT 246
Qy      371 MetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIlePro 390
Db      247 TACAACCTTGAATGCAATCCAGATGACTATGAGGAGTGAATAAATTCACATCCCTCG 306
Qy      391 HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410
Db      307 GAAAGATCTGGCGCGGAGCTGCTCTCTATTAACAACGAGAGCGGCGACTTTGCCATT 366
Qy      411 ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGly 430
Db      367 GTCAAAATTCACCAAGTGTCTCTGGACTACACCGGCCACATCCTGGACACCGCCGCC 426
Qy      431 IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys 450
Db      427 ATCTTTAAAGACTACTGTGAGATCATTTGCTACTCCTTCCCTTCGTGAGCAGAACTGC 486
Qy      451 GluMetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAsp 470
Db      487 AGCATGAAGCTGGGCACCTGGACCTATGACGGCTCTGTGTGGCCATAACCCGAAAGT 546
Qy      471 GluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuGlyVal 490
Db      547 GACCAG---CCGACCTGAGTAACCTCATGGAGCGGGAGTGGGTGATCAAGAAAGCT 603
Qy      491 ProGlyLysArgAsnGluIleTyrTyrAsnCysCysPro---GluProTyrIleAspIle 509
Db      604 CGGGCTGGAAGCAGCACTGGTGTCTACTCTCTGCTGCCACCACCTCCTACTCGACATC 663
Qy      510 ThrPheAlaIleIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIlePro 529
Db      664 ACCTACCACCTCGCTATGACGCGCTCCCTCTACTTTCATTGTCAACGCTCATCATCCC 723
Qy      530 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 549
Db      724 TGCTGTCTTCTCTCTTCTTAACACGCTGGTGTCTACTCTGCCACGACTCAGGGAG 783
Qy      550 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 569
Db      784 AAGATGACGCTGAGCATCTCTGTCTTACTGTCTGCTGACCGTGTCTCTTCTGTGTGTG 843
Qy      570 GluThrMetProAlaThrSerAspAlaValProLeu----- 581
Db      844 GAGTAATCCCTTCCACCTCCAGCGCTGCCCTCATGGGAAGTATATGTGTTCACC 903
Qy      581 ----- 581
Db      904 ATGGTCTTGTTCATTGGTCCATCATCACCGTCTATCGTCAACACACACCCCGT 963
Qy      582 -----TrpIleArgIleValPheLeuCysTrpLeuPro 592
Db      964 TCGCCCGAGCACCACATCATGCGCGAGTGGTGGGAGGTTTATCGACACATATCCCA 1023
Qy      593 TrpIleLeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrPro 612
Db      1024 AACATCATGTGTTTCTCCCAATGAAGAAGACCA----- 1056
Qy      613 CysSerAspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGlu 632
Db      1057 -----TCCAGAGAT 1065
Qy      633 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsn 652

```

```

...      |||      ...      :|||:|||||      |||
Db 1066 AAACAGAGAAAGATTTTACAGAGACATAGATATCTGAC----- 1110

Qy 653 CysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyArgThrVal 672

Db 1110 ----- 1110

Qy 673 TyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAla 692

Db 1111 -----ATCTCTGGGAACCGCGTCTCT-----CCACCTATG 1140

Qy 693 ValThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeuHis 712

Db 1141 GGCTTTCATCTCGCTGATCAAGCACCT-----GAGTCAAAAGCGCCATCGAG 1191

Qy 713 GluIleArgPheIleThrAspGlnLeuArgLysAspGluCysAsnAspIleAlaAsn 732

Db 1192 GGCGTGAAGTACATTGCAGAGACCATGAAGTCAGACGAGGAGTCCATTAACGCCGCTGAG 1251

Qy 733 AspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPhe 752

Db 1252 GAATGGAAGTATGTTGCTCATGATGATGATCATCTCCCGGAGTCTTTATGCTGTG 1311

Qy 753 AlaIleLeuAlaThrIleAlaVal 760

Db 1312 TGCTCATCGGACGCTGCTGTG 1335

AK029177      4290 bp      mRNA      linear      HTC 18-SEP-2003
Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
library, clone:4831406G09 product:cholinergic receptor, nicotinic,
alpha polypeptide 1 (muscle), full insert sequence.
AK029177
AK029177.1 GI:26325165
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, I., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

```

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

polyA_signal

polyA_site

ORIGIN

Alignment Scores:

Pred. No.: 4,93e-60 Length: 4290

Score: 797.00 Matches: 169

Percent Similarity: 51.97% Conservative: 95

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 4290)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohmoto, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1..4290

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:4831406G09"

/db_xref="MGI:2391229"

/db_xref="taxon:10090"

/clone="4831406G09"

/tissue type="head"

/clone lib="RIKEN full-length enriched mouse cDNA library"

/dev stage="0 day neonate"

21..1394

/note="unnamed protein product; cholinergic receptor,

nicotinic, alpha polypeptide 1 (muscle) [MGI:87885,

GB|NM_007389, evidence: BLASTN, 99%, match=1763]

putative"

/codon start=1

/protein_id="BAC26337.1"

/db_xref="GI:26325162"

/translation="MELSTVLLLLCLCSAGLVGSEHETRLVAKLFEDYSSVVRVED

HREIVQVTVGLQIQLINVDENVQIVTNRLLKQWVDYNLKNWDDDDGGVKKIHP

EKIWRPDLVYNNADGDAIYVKFTVLDYDGHITWTPAIFPSYCEIIVTHFPDEQ

NCMKLGVTYDGSVAINPESQDPLSNFMESEGVKEARQGWKHWVYSCCTTPY

LDITVHFVMDRLPLFVIVNVIICLLFSLFVFLPTDGERKTLSTISLLSTVF

LLVIVLPTSSAVPLIGKMLMTWVFVIAIITIVIVNTHHSPSTHMPWVRK

VFDITPIMFSTWKRPSRQKREITDIDISGKPGPPGPHSPILIKHPEV

LSAIGVKYIATMTKSDQESNNAEENKIVAMVMDHILLGVFMVLCVLTGLAVFAGRL

1ELHQQG"

4267..4272

/note="putative"

4290

/note="putative"

Best Local Similarity: 33.27% Mismatches: 140
 Query Match: 19.71% Indels: 104
 DB: 11 Gaps: 8

US-09-303-232-2 (1-770) x AK029177 (1-4290)

QY 291 SerTrpIlePheLeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr 310
 DB 30 TCGATGTTCTCTGCTAGGCTC---TGCTCCGCTGGCTTGTCTGGGCTCCGAA 86
 QY 311 HisGlnLysArgLeuLeuHisAspLeuAspProTyrAsnThrLeuGluArgProVal 330
 DB 87 CATGAGCGCTCTGGTGGCAAGCTCTTTGAAGACTACAGCAGTAGTCGGCCAGTG 146
 QY 331 LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAsp 350
 DB 147 GAGACACCGCTGAGATTGTACAGTACCCGTGGTCTACAGCTGATCCAGCTTATCAAT 206
 QY 351 ValAspGlnLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp 370
 DB 207 GTGATGAAGTAATCAGATTGTGACACCAATGTACGTCTGAACACCAATGGTGCAT 266
 QY 371 MetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProPro 390
 DB 267 TACAACCTTGAATGGAATCCAGATGACTATGGAGAGTGAATAAATTTCATCCCTCG 326
 QY 391 HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410
 DB 327 GAAAGATCTGGCGCGGACGTCGTCTCTATACCAACGACAGCGGCACTTTTGGCATT 386
 QY 411 ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGly 430
 DB 387 GTCAATTCACCAAGTGCTCCTGGATACACCGGCCACATCACCTGGACACCCGAGCC 446
 QY 431 IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys 450
 DB 447 ATCTTTAAAAGCTACTGTGAGATCATTTGTCTACTCACCTTCGATGACGAACTGC 506
 QY 451 GluMetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnAsp 470
 DB 507 AGCATGAGCTGGCAGCTGGACCTATGACGGCTCTGTGTGGCCATTAAACCGGAAAGT 566
 QY 471 GluThrGlyGlyAspIleSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyVal 490
 DB 567 GACCAG---CCCGACCTGAGTAATCTCATGAGAGCGGGAGTGGGTGATCAAGGAAGCT 623
 QY 491 ProGlyLysArgAsnGlnLysTyrTyrAsnCysPro---GluProTyrIleAspIle 509
 DB 624 CGGGCTGGAAGCACTGGGTGTCTACTCTGTCTGCCCCACCACTCCCTACCTGGACATC 683
 QY 510 ThrPheAlaIleIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIlePro 529
 DB 684 ACTACCACTTGTGTATGAGCGCTCGCCCTCTACTTATTTGTCACGTCATCATTCCTCC 743
 QY 530 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 549
 DB 744 TGCCTGCTCTCTCTCTTAACCAAGCTGGTGTCTACTCGCCACAGACTCAGGGGAG 803
 QY 550 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 569
 DB 804 AAGATGACGCTGAGCATCTGTGTTACTGTCTCCTGACCGTGTCTCTCTGGTCAATTGTG 863
 QY 570 GluThrMetProAlaThrSerAspAlaValProLeu----- 581
 DB 864 GAGCTAATCCCTTCCACCTCCAGCGCTGCCCCCTGATCGGGAAGTATATGTTTCAACC 923
 QY 581 ----- 581
 DB 924 ATGGTCTTTGTCATTGGTCCATCATCATCACCGTCATCGTCAATCAACACACACACCGCT 983
 QY 582 -----TrpIleArgIleValPheLeuCysTrpLeuPro 592
 DB 984 TCGCCGAGCACCCACATCATGCGGAGTGGTGGGAGGTTTATTCGACACATATCCCA 1043

QY 593 TrpIleLeuArgMetSerArgProGlyArgProLeuLeuLeuGluPheProThrThrPro 612
 DB 1044 AACATCATGTTTCTCCACAATGAAAGACCA-----TCCAGAGAT 1085
 QY 613 CysSerAspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGlu 632
 DB 1077 ----- 1077
 QY 633 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsn 652
 DB 1086 AAACAAGAGAAAAGATTTTACAGACACATAGATATATCTGAC----- 1130
 QY 653 CysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrVal 672
 DB 1130 ----- 1130
 QY 673 TyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAla 692
 DB 1131 -----AFTCTGGGAAGCCGGTCTCT-----CCACCTATG 1160
 QY 693 ValThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLys 712
 DB 1161 GGCTTTCACTCTCCGTGATCAAGCACCT-----GAGGTGAAAGCGCGCTGAG 1211
 QY 713 GluIleArgPheIleThrAspGlnLeuArgLysAspGluCysAsnAspIleAlaAsn 732
 DB 1212 GGGTGAAGTACATGTCAGAGACCATGATGATCAGACGAGTCCATATACCGCGCTGAG 1271
 QY 733 AspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPhe 752
 DB 1272 GAATGGAAGTATGTGCGATGGTATGATGCATCATCTCTCGAGTCTTTATGCTGGTG 1331
 QY 753 AlaIleLeuAlaThrIleAlaVal 760
 DB 1332 TGCTCATCGGAGCGCTGGCTGIG 1355

RESULT 15
 AY402877
 LOCUS
 DEFINITION Pan troglodytes CHRN4 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY402877
 VERSION AY402877.1 GI:39758860
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 REFERENCE 1 (bases 1 to 1442)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1442)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 1..1442
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"

```
Gene      <1. >1442
           /gene="CHRNA4"
           /locus_tag="HCM1370"

ORIGIN

Alignment Scores:      Length: 1442
Pred. No.:             3,22e-60
Score:                 791.00
Percent Similarity:    51.29%
Best Local Similarity: 35.05%
Query Match:           19.56%
DB:                    29
Gaps:                  13

US-09-303-232-2 (1-770) x AY402877 (1-1442)

Qy 306 CysLeuAlaGlyTyrHisGluLeuLeuHisAspLeuLeuLeuPro-----Tyr 323
Db 6 TGGCGGTGGCCAAATGGCGAGGAAAAGCTGATGAGCAGACCTTCTGAACAAAACCGTTAC 65
Qy 324 AsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeu 343
Db 66 AATAACCTGATCCGCCAGCCAGCCAGCTCCTCAGCTCATCTCTCAATCAAGCTGAGTGC 125
Qy 344 ThrLeuMetGlnIleLeuAspValAspGluLysAsnGlnLeuValThrAsnValTrp 363
Db 126 TCCCTGGCCCACTATCAGCGTGAATGAGCGAGAGCAGATCATGACCACCAATGTCTGG 185
Qy 364 LeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyrGlyVal 383
Db 186 CTGAACAGGAATGAGCTGATGATCCGCTGACCTGGAACAGCTCCCGTACGAGGGTGTG 245
Qy 384 LysAspLeuArgIleProHisArgIleTrpLysProAspValLeuMetTyrAsnSer 403
Db 246 AACATCCGTGAGATCCCTCGAAAGCGCACTCTGGTTCCTGACATCGCTGCTTACCAAC 305
Qy 404 AlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValAlaArgAsnGlySer 423
Db 306 GCCGACGGGACTATGAGGTCTGTCTACACCACTTGTATGTCGGTCCCAATGGCAGC 365
Qy 424 CysLeuTrpValProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPhe 443
Db 366 CTCCTGTGGTCCCGCTGCCATCTACAGAGCGCTGCAAGATTGAGGTGAAGTACTTT 425
Qy 444 ProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAspGlyPheGln 463
Db 426 CCGTTCGACAGCAAGACTGACCCCTCAAGTTCCCTCTCGGACCTATGACCACAGCGAG 485
Qy 464 LeuAspLeuGlnLeuAspGluThrGlyGlyAspIleSerSerTyrValLeuAsnGly 483
Db 486 ATAGACATGGTCTCTCATGAGGCCACA--GCCAGCATGGATGACTTTACTCCCATGGT 542
Qy 484 GluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysPro 503
Db 543 GAGTGGGACATAGTGGCCCTCCAGGGAGAGACA-----GTGAACCCACAGGAC 593
Qy 504 GluProTyrIleAspIleThrPheAlaIleIleArgArgArgThrLeuTyrTrpPhe 523
Db 594 CCCAGCTACGGAGCTGACTTACGACTTCATCATCAGCGCAAGCCCTGTGTCTACACC 653
Qy 524 PheAsnLeuIleIleProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeu 543
Db 654 ATCAACCTCATATCCCTCGTGTCTACCACTGTGTGGCCATCCTTGTCTTCTACCTG 713
Qy 544 ProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrVal 563
Db 714 CCATCCAACTGGCGGAGAGATGACACTGTGCATCTCAGTGTGCTGCGCATTTACATTC 773
Qy 564 PheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeu----- 581
Db 774 TTCCTGCTGCTCATCTCCANGTNGGCGCACCCACCTCCCTCGATGTGCTCTCATCGGC 833
Qy 581 ----- 581
```

Search completed: May 9, 2004, 04:19:59
Job time : 5420.49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 12:59:49 ; Search time 5894.18 Seconds
(without alignments)
10942.062 Million cell updates/sec

Title: US-09-303-232-3_COPY_335_1822
Perfect score: 1488
Sequence: 1 atggggcggggcgccgcgcg.....cgccacacatcatgtgtgcg 1488

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1488	100.0	3629	3	AF143846	AF143846 Heliothis
2	1488	100.0	3700	6	AX009612	AX009612 Sequence
3	1488	100.0	3701	6	E58347	E58347 Nucleic aci
4	615	41.3	2023	3	AF321446	AF321446 Drosophil
5	611.8	41.1	1665	3	DME554209	AF554209 Drosophil
6	611.8	41.1	1699	3	BT011147	BT011147 Drosophil
7	611.8	41.1	2023	3	AF321447	AF321447 Drosophil
8	608.6	40.9	2023	3	AF321445	AF321445 Drosophil
9	583.8	39.2	2907	3	AF272778	AF272778 Drosophil
10	575	38.6	3029	3	AF143847	AF143847 Heliothis
11	575	38.6	3109	6	E58348	E58348 Nucleic aci
12	575	38.6	3109	6	AX009614	AX009614 Sequence
13	568.4	38.2	1683	3	DME554210	AJ554210 Drosophil
14	553.6	37.2	2068	3	AF321448	AF321448 Drosophil
15	519.6	34.9	2110	3	AF321449	AF321449 Drosophil
16	512.8	34.5	2886	6	E58346	E58346 Nucleic aci
17	512.8	34.5	2886	6	AX009610	AX009610 Sequence
18	455.4	30.6	2834	3	AY036613	AY036613 Drosophil
19	410.6	27.6	1848	10	MUSNARS	L37663 Mus musculu
20	410	27.6	2107	9	AF486623	AF486623 Macaca mu
21	407.4	27.4	2088	10	AF225980	AF225980 Mus muscu
22	405.6	27.3	1509	6	AX054567	AX054567 Sequence
23	405.6	27.3	1876	6	AR282833	AR282833 Sequence
24	405.6	27.3	1876	6	AX719088	AX719088 Sequence
25	405.6	27.3	1876	9	HSU62436	U62436 Human nicot
26	405.2	27.2	1876	6	AR055255	AR055255 Sequence
27	405.2	27.2	1876	6	AR071403	AR071403 Sequence
28	405.2	27.2	1876	6	AR173187	AR173187 Sequence
29	405.2	27.2	1876	6	AR224030	AR224030 Sequence
30	404	27.2	1509	6	AX054575	AX054575 Sequence
31	404	27.2	1509	6	AX054577	AX054577 Sequence
32	404	27.2	1509	9	HSNACHRA7	Y08420 H.sapiens m
33	404	27.2	3030	10	S53987	S53987 nicotinic r
34	402.4	27.0	1509	6	AX054579	AX054579 Sequence
35	401.8	27.0	2106	10	RATNARAD	I31619 Rattus ratt
36	400.8	26.9	1916	9	AF385585	AF385585 Homo sapi
37	400.8	26.9	1977	9	HSU40583	U40583 Human alpha
38	400.8	26.9	2087	9	HSARA7A	X70297 H.sapiens m
39	397.6	26.7	1555	9	HSCHRNA7A	Z23141 H.sapiens C
40	396.4	26.6	1551	4	BT47	X93604 B.taurus mR
41	396	26.6	1559	9	HUMATNAR	L25827 Human a7 ni
42	394	26.5	1590	6	AR261850	AR261850 Sequence
43	394	26.5	1590	6	BD023656	BD023656 Variant h
44	364	24.5	1668	5	AY298752	AY298752 Takifugu
45	357.6	24.0	1542	5	AY299466	AY299466 Takifugu

ALIGNMENTS

RESULT 1
LOCUS AF143846
DEFINITION Heliothis virescens putative nicotinic acetylcholine receptor alpha
7-1 subunit mRNA, complete cds.
ACCESSION AF143846
VERSION AF143846.1 GI:4895004
KEYWORDS Heliothis virescens (tobacco budworm)
SOURCE Heliothis virescens
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3629)
AUTHORS Schulte,T., Oellers,N. and Adamczewski,M.

TITLE Putative alpha subunits of insect nicotinic acetylcholine receptors more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than to other insect nicotinic acetylcholine receptor alpha subunits

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3629)
AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
TITLE Direct Submission
JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18., Leverkusen 51368, Germany
FEATURES Location/Qualifiers
source 1..3629
/organism="Heliothis virescens"
/mol_type="mRNA"
/db_xref="taxon:7102"
335..1825
/note="hvnachra7-1"
/codon_start=1
/product="putative nicotinic acetylcholine receptor alpha 7-1 subunit"
/protein_id="AAD32697.1"
/db_xref="GI:4895005"
/translation="MGGRRSHLAAPAGLLLLCLLWPRGARCYHEKRLHLLHLDH YNVLPRVNESDPLQSLFGLTLMQI IDVDEKNQLLNIWLKLEWMDNLRWNTSDF GGVKDLRVPRLMKFSGWTDYGLDLOLEGGDDISSFTVINGEWELIGVPGKRNK LIYNCCPEPYDITFAVIRKTLIYFENLIVPCVLJASMLLGLFTLPDPSGKLSLG VTLLSLVFLDMVAETMPATSDAVPLLLGTTFNCIMENVASSVSTILLNIYHRHAD THSDMTDRCVFLVLPWLWRMSRPGSATPPPPPPDLERLSKSLLLANVLIDDDFRHPQAOQPCYRGRGEENAGLAHSGFVDYELSLKLEIRVTDQWRKD DEDADISRDWKFAMVVDRLCLLIPTLTIATLAVLLSAPHIMVS"

CDS
Query Match 100.0%; Score 1488; DB 3; Length 3629;
Best Local Similarity 100.0%; Pred. No. 7.2e-291;
Matches 1488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCGCGGGCGCGCGCTCGACATTTGGCGGGCGCGCGGGCGCTGCTGCTGCTG 60
DB	335	ATGGCGGGGGCGCGCGCTCGACATTTGGCGGGCGCGCGGGCGCTGCTGCTGCTG 394
QY	61	TGCTGCTGCTGCGCGGGGGCGCGCTGCGGGTACCAAGAGCGGCTACTGCACAC 120
DB	395	TGCTGCTGCTGCGCGGGGGCGCGCTGCGGGTACCAAGAGCGGCTACTGCACAC 454
QY	121	CTATTGGACCACTCAACGCTACTGAGAGGCGCGCTGCTGCGGGTACCAAGAGCG 180
DB	455	CTATTGGACCACTCAACGCTACTGAGAGGCGCGCTGCTGCGGGTACCAAGAGCG 514
QY	181	CTCTCTCTGCGCTCAACGCTATGAGATCATCGACGCTGAGAGGAGAACAGCTTT 240
DB	515	CTCTCTCTGCGCTCAACGCTATGAGATCATCGACGCTGAGAGGAGAACAGCTTT 574
QY	241	ATPACAAACATCTGCTAAACATAGAGTGAATGATGAACTTGAAGTGAACACTTCA 300
DB	575	ATPACAAACATCTGCTAAACATAGAGTGAATGATGAACTTGAAGTGAACACTTCA 634
QY	301	GATTTGCGGGGTCAAGATTTAAGTGTGACACCCACAGACTATGGAACACGACGTC 360
DB	635	GATTTGCGGGGTCAAGATTTAAGTGTGACACCCACAGACTATGGAACACGACGTC 694
QY	361	CTTATGTACACAGCGCGGACGAAGGTTTCGACAGCAGCTATCCAAACGACGTTGG 420
DB	695	CTTATGTACACAGCGCGGACGAAGGTTTCGACAGCAGCTATCCAAACGACGTTGG 754
QY	421	CGGAACAAACGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB	755	CGGAACAAACGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 814
QY	481	GACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB	815	GACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874

ORIGIN

QY	541	TATGATGTTTATCAGTTGGATCTACAACTACAGGATGAAGGGGGGAGATATAGCAGT 600
DB	875	TATGATGTTTATCAGTTGGATCTACAACTACAGGATGAAGGGGGGAGATATAGCAGT 934
QY	601	TTTGTACGAATGGCGAATGGAGTTTATAGGAGTCCCGGCAAGCGCAAGATCTAC 660
DB	935	TTTGTACGAATGGCGAATGGAGTTTATAGGAGTCCCGGCAAGCGCAAGATCTAC 994
QY	661	TACAACTGTTTGTCCGGAGCCATACATCGACATCACAGTTTGGCGTGGTGCATCCGAGGAAA 720
DB	995	TACAACTGTTTGTCCGGAGCCATACATCGACATCACAGTTTGGCGTGGTGCATCCGAGGAAA 1054
QY	721	ACGCTCTACTACTTCTTCAATCTCATCGTGCCTCGCTGCTCATCGCTCCATGCTCTA 780
DB	1055	ACGCTCTACTACTTCTTCAATCTCATCGTGCCTCGCTGCTCATCGCTCCATGCTCTA 1114
QY	781	TTGGGGTTACCTTGCCTCCAGACTCCGGAGAAAAGTTGCTTTAGGTGTGACGATATTA 840
DB	1115	TTGGGGTTACCTTGCCTCCAGACTCCGGAGAAAAGTTGCTTTAGGTGTGACGATATTA 1174
QY	841	CTGTGCTGACGCTGTTCTCTCAACATGTTGGCGGAGACGATGCCAGCGACGCTCGACGCC 900
DB	1175	CTGTGCTGACGCTGTTCTCTCAACATGTTGGCGGAGACGATGCCAGCGACGCTCGACGCC 1234
QY	901	GTGCCCTTGTCCGCACTACTTCAACTGATCATGTTTATGTTGCTGCTGCTGCTGCTGCT 960
DB	1235	GTGCCCTTGTCCGCACTACTTCAACTGATCATGTTTATGTTGCTGCTGCTGCTGCTGCTGCT 1294
QY	961	TCCACCATACTGATCTCTCAACTACACACCGGAGCGAGACACTCAGGAAATGATGAT 1020
DB	1295	TCCACCATACTGATCTCTCAACTACACACCGGAGCGAGACACTCAGGAAATGATGAT 1354
QY	1021	TGGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB	1355	TGGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1414
QY	1081	TCGGGACGACG 1140
DB	1415	TCGGGACGACG 1474
QY	1141	CGCTCTCAAGTCTGCTAGCGAAGTCTGACATCGATGACGCTTCCGCGCACCGG 1200
DB	1475	CGCTCTCAAGTCTGCTAGCGAAGTCTGACATCGATGACGCTTCCGCGCACCGG 1534
QY	1201	CAAGCGACGACG 1260
DB	1535	CAAGCGACGACG 1594
QY	1261	TTGGCGGCGACAGTTGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB	1595	TTGGCGGCGACAGTTGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1654
QY	1321	AGAGTCATCAGATCAGATCGGAGGACGAGAGATCGGAGATTTCCGCGGACGCTG 1380
DB	1655	AGAGTCATCAGATCAGATCGGAGGACGAGAGATCGGAGATTTCCGCGGACGCTG 1714
QY	1381	AAGTTTCGCGCGCATGCTGCTGCGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB	1715	AAGTTTCGCGCGCATGCTGCTGCGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1774
QY	1441	ATCGCGACGCTAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1488
DB	1775	ATCGCGACGCTAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1822

RESULT 2
AX009612
LOCUS
DEFINITION Sequence 3 from Patent EP0962528.
ACCESSION AX009612
VERSION AX009612.1 GI:9996844
KEYWORDS Heliothis virescens (tobacco budworm)

ORGANISM	Heliothis virescens	Db	875	TATGATGGTTATCAGTTGGATCTCAACTACAGATGAAGGGGGGAGATATAAGCAGT	934
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.	Qy	601	TTTGTACGAATGGGAATGGAGTTAATAGGAGTCCCGGCAAGCCACAGATCTAC	660
REFERENCE	1	Db	935	TTTGTACGAATGGGAATGGAGTTAATAGGAGTCCCGGCAAGCCACAGATCTAC	994
AUTHORS	Adamczewski, M.D., Schulte, J.D. and Oellers, N.D.	Qy	661	TACAACTGTGTCCGGAGCCATACATGACATCACGTTTCGGTGGTGGATCCGGAGAAA	720
TITLE	Nucleic acids encoding acetylcholin-receptor subunits from insects	Db	995	TACAACGTGTTCGGAGCCATACATGACATCACGTTTCGGTGGTGGATCCGGAGAAA	1054
JOURNAL	Patent: EP 0962528-A 3 08-DEC-1999;	Qy	721	ACGCTCTACTACTTCTTCAATCTGATCGTCCCTGGTGTCTCATCGCTCCATGGCTCTA	780
FEATURES	BAYER AG (DE)	Db	1055	ACGCTCTACTACTTCTTCAATCTGATCGTCCCTGGTGTCTCATCGCTCCATGGCTCTA	1114
source	Location/Qualifiers	Qy	781	TTGGGGTTCACTTGGCTCCAGACTCCGGAGAAAAGTTGTTTGTAGTGTGACATATTA	840
	1. 3700	Db	1115	TTGGGGTTCACTTGGCTCCAGACTCCGGAGAAAAGTTGTTTGTAGTGTGACATATTA	1174
CDS	/organism="Heliothis virescens"	Qy	841	CTGTGTTGACGGTGTTCCTCAACATGTCGGCGAGACGATCCAGCGAGCTCGACGCC	900
	/mol_type="unassigned DNA"	Db	1175	CTGTGTTGACGGTGTTCCTCAACATGTCGGCGAGACGATCCAGCGAGCTCGACGCC	1234
	/db_xref="taxon:7102"	Qy	901	GTGCGCTTGTCTCGGACACTTCAACTGCATCATGTTCAATGTTGGTGGCTTCTCCGTCGC	960
	335..1825	Db	1235	GTGCGCTTGTCTCGGACACTTCAACTGCATCATGTTCAATGTTGGTGGCTTCTCCGTCGC	1294
	/note="unnamed protein product"	Qy	961	TCACACATGATGCTTCACTACACACCGGACGACACACTACAGAAATGAGTGAT	1020
	/codon_start=1	Db	1295	TCCACCATGATGCTTCACTACACACCGGACGACACACTACAGAAATGAGTGAT	1354
	/protein_id="CAC07500.1"	Qy	1021	TGATTCGTTGCGGTGTTCTTATTTGGTCTGCGGTGGTGTGCGCATGTACGCGCCCGGC	1080
	/db_xref="GI:9996845"	Db	1355	TGATTCGTTGCGGTGTTCTTATTTGGTCTGCGGTGGTGTGCGCATGTACGCGCCCGGC	1414
	/translation="MGRARRSHLAAPAGLLLLCLLWPRGARGVHEKRLHLHLDH YNYLRFVNESPLQSLFGLTLMQIIDVDEKNQLLITNLKLEWNNLRNWTSD GGVKDLRVPRLHMKPDLVLMYNGADFSTPTNVVNRNGSLYVPPGIFKSTCKI DITWFDDORCEMKGSTYDGVOLDLQDEGGDISSFVTNGEWELIGVPEKNE IYNCCPEYIDITFAVIRKTLXYFNLI VPCVLASMLLGLFTLPDPSGEKLSIG VTLLSLTVFLNVAETMPATSDAVPLLTGFYFNCIMFWASSVTVTLLINLHHRAD THMSDMIRCVFLWLPVLRMSRPGSATTPPARVPPPPDLERRSKSLLANVLD IDDFRFPQAQPCQCYRGGGEAGLAHSCFGVDYBELSLILKEIRVITDQMRKD DEDADISRDKFAAMVVDRLIIFTLITTIATLAVLLSAPHIMVS"	Qy	1081	TCGGCGACGACGCGCGCGCGGTACTCTCCGCGCGGACCTCGGAGCTCGCGAG	1140
ORIGIN		Db	1415	TCGGCGACGACGCGCGCGCGGTACTCTCCGCGCGGACCTCGGAGCTCGCGAG	1474
Query Match	100.0%; Score 1488; DB 6; Length 3700;	Qy	1141	CGCTCTTCAAGTGTCTCTAGCGAAGTGTCTGACATCATGACGACTTTCGCGACCCG	1200
Best local similarity	100.0%; Pred. No. 7.2e-291;	Db	1475	CGCTCTTCAAGTGTCTCTAGCGAAGTGTCTGACATCATGACGACTTTCGCGACCCG	1534
Matches 1488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Qy	1201	CAAGCGACGACGCGCAATGCTGCCGATCTACAGGGGGGTGAGGAATGGCGCGGG	1260
1	ATGGCGGGCGGGCGCGCTCGCACTTGGCGCGCGCGCGCGCTGCTGCTGCTG 60	Db	1535	CAAGCGACGACGCGCAATGCTGCCGATCTACAGGGGGGTGAGGAATGGCGCGGG	1594
335	ATGGCGGGCGGGCGCGCTCGCACTTGGCGGGCGCGCGCGCTGCTGCTGCTG 394	Qy	1261	TTGCGCGCGACAGTTGCTTGGTGTGACGTACAGGTCTCCCTCATCTTCTGAAGAGATT	1320
61	TGCTGCTCTGCGCGAGGGGGCGCGCTGCGGGTACCAAGAGAGCGGCTACTCCACC 120	Db	1595	TTGCGCGCGACAGTTGCTTGGTGTGACGTACAGGTCTCCCTCATCTTCTGAAGAGATT	1654
395	TGCTGCTCTGCGCGAGGGGGCGCGCTGCGGGTACCAAGAGAGCGGCTACTCCACC 454	Qy	1321	AGAGTCATCAGATCAGATCGCGAAGACGACGACGATTCGGGAGCTGG 1380	
121	CTATTGACCACTACAGCTACTCGAGAGCGCGCTGCTCAACGAGAGCGCGCTGCGAG 180	Db	1655	AGAGTCATCAGATCAGATCGCGAAGACGACGACGATTCGGGAGCTGG 1714	
455	CTATTGACCACTACAGCTACTCGAGAGCGCGCTGCTCAACGAGAGCGCGCTGCGAG 514	Qy	1381	AAGTTCGCGCGCATGCTGCGGACAGACTGTGCTTATTTATCTTACCTGTTTCAACATC 1440	
181	CTCTCTCTGCGCTCAGCTCATGACATCATCGACGTGGAGAGAACACGCTTTTA 240	Db	1715	AAGTTCGCGCGCATGCTGCGGACAGACTGTGCTTATTTATCTTACCTGTTTCAACATC 1774	
515	CTCTCTCTGCGCTCAGCTCATGACATCATCGACGTGGAGAGAACACGCTTTTA 574	Qy	1441	ATCGCCACGCTAGCGGTGCTGTCTCGCGCCACACATCATGTGTGTG 1488	
241	ATAACAAACATCTGGCTAAACCTAGTGGAAATGATGAACCTTGAGTGGAACACTTCA 300	Db	1775	ATCGCCACGCTAGCGGTGCTGTCTCGCGCCACACATCATGTGTGTG 1822	
575	ATAACAAACATCTGGCTAAACCTAGTGGAAATGATGAACCTTGAGTGGAACACTTCA 634	LOCUS	ES8347	3701 bp DNA linear PAT 18-JUN-2001	
301	GATTTGCGGGGTCAAGATTAAAGTGGCCACCCACAGACTATGGAACACGACGTC 360	DEFINITION	Nucleic acid encoding insect actyl choline receptor subunit.		
635	GATTTGCGGGGTCAAGATTAAAGTGGCCACCCACAGACTATGGAACACGACGTC 694	ACCESSION	ES8347		
361	CTTATGTACACAGCGCGAGAGGGTTGACAGACGATTCACAAAGAGTGGTGTG 420	VERSION	ES8347.1	GI:13019346	
695	CTTATGTACACAGCGCGAGAGGGTTGACAGACGATTCACAAAGAGTGGTGTG 754	KEYWORDS	JP 2000023680-A/2.		
421	CGGAACACGCGCTCGTGTCTGCTAGCTGCGCGCGCGGCACTTCAAGAGCACCTGCAAGATC 480	SOURCE	Heliothis virescens (tobacco budworm)		
755	CGGAACACGCGCTCGTGTCTGCTAGCTGCGCGCGCGGCACTTCAAGAGCACCTGCAAGATC 814	ORGANISM	Heliothis virescens		
481	GACATCACTGTTCCCTTTCGACGACCAACGATCGAGATGAAGTTTGGCAGCTGGACT 540		Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;		
815	GACATCACTGTTCCCTTTCGACGACCAACGATCGAGATGAAGTTTGGCAGCTGGACT 874				
541	TATGATGGTTATCAGTTGGATCTACAACTACAGGATGAAGGGGGCGGAGATATAAGCAGT 600				

Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.	
REFERENCE	1 (bases 1 to 3701)
AUTHORS	Martin,A., Nadja,E. and Thomas,S.
TITLE	Nucleic acid encoding insect acetyl choline receptor subunit
JOURNAL	BAYER AG
COMMENT	
OS	Heliothis virescens
PN	JP 2000023680-A/2
PD	25-JAN-2000
PF	26-APR-1999 JP 1999118159
PR	04-MAY-1998 DE 19819829.9
PI	MARTIN ADAMUTSUEBUSHI,NADJA ERASU,THOMAS SCHULTE PC
C12N15/09,A01K67/033,C07K14/705,C07K16/28,C12N1/21,C12N5/10, PC	
C1201/68,	
PC	G01N33/15,G01N33/50/(C12N1/21,C12R1/19),C12N15/00,C12N5/00 CC
FEATURES	
source	
ORIGIN	
Query Match 100.0%; Score 1488; DB 6; Length 3701;	
Best Local Similarity 100.0%; Pred. No. 7.2e-291; Indels 0; Gaps 0;	
Matches 1488; Conservative 0; Mismatches 0;	
QY	1 ATGGGCGGGCGGCGCGCTCGACATTGGCGGCGCGCGGCGCTGCTGCTGCTG 60
DB	335 ATGGGCGGGCGGCGCGCTCGACATTGGCGGCGCGCGGCGCTGCTGCTGCTG 394
QY	61 TGCTGCTGCTGCGGAGGGGGGCGCTGCGGGGTACACGAGAGCGGCTACTGCAAC 120
DB	395 TGCTGCTGCTGCGGAGGGGGGCGCTGCGGGGTACACGAGAGCGGCTACTGCAAC 454
QY	121 CTATTGGACCACTCAACGTTACTGGAGAGGCGGCTGCTCAACGAGAGCGGCGTGCAG 180
DB	455 CTATTGGACCACTCAACGTTACTGGAGAGGCGGCTGCTCAACGAGAGCGGCGTGCAG 514
QY	181 CTCTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB	515 CTCTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574
QY	241 ATACAAACATCTGGCTAAACTAGAGTGGATGATGATGATGATGATGATGATGAT 300
DB	575 ATACAAACATCTGGCTAAACTAGAGTGGATGATGATGATGATGATGATGATGAT 634
QY	301 GATTTCCGGCGGGTCAAGATTTTAAAGTGGCCACCCACAGACTATGGAACCCAGAGTC 360
DB	635 GATTTCCGGCGGGTCAAGATTTTAAAGTGGCCACCCACAGACTATGGAACCCAGAGTC 694
QY	361 CTTATGTACAAACAGCGGAGCAAGGTTTCAGACAGAGTATCCAAACGAACTGGTGTG 420
DB	695 CTTATGTACAAACAGCGGAGCAAGGTTTCAGACAGAGTATCCAAACGAACTGGTGTG 754
QY	421 CGGAACAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB	755 CGGAACAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
QY	481 GACATCACTGGTTCCCTTTCGACGACCAACGATGCGAGATGAAGTTTGGCAGCTGGACT 540
DB	815 GACATCACTGGTTCCCTTTCGACGACCAACGATGCGAGATGAAGTTTGGCAGCTGGACT 874
QY	541 TATGATGGTTATCAGTTGGATCTCACTACAGATGAAGGGGCGGAGATATTAAGCAGT 600
DB	875 TATGATGGTTATCAGTTGGATCTCACTACAGATGAAGGGGCGGAGATATTAAGCAGT 934
QY	601 TTTTGTACGAATGGCAATGGAGTTAATAGAGTCCCGGCAAGCGCAAGCATCTAC 660
DB	935 TTTTGTACGAATGGCAATGGAGTTAATAGAGTCCCGGCAAGCGCAAGCATCTAC 994

QY	661 TACAACTGTTCCCGAGCCATACATCGACATCAGTTTCCGCTGGTGTATCCGAGGAAA 720
DB	995 TACAACTGTTTCCCGAGCCATACATCGACATCAGTTTCCGCTGGTGTATCCGAGGAAA 1054
QY	721 ACGCTCTACTACTCTTCAATCTGATCGTCCCTCGCTGCTCATCCCTCCATGCTCTA 780
DB	1055 ACGCTCTACTACTCTTCAATCTGATCGTCCCTCGCTGCTCATCCCTCCATGCTCTA 1114
QY	781 TTGGGTTTCACTTCCCTCCAGACTCCGAGAGAAAAGTTGTCTTTAGGTGTGACGATATTA 840
DB	1115 TTGGGTTTCACTTCCCTCCAGACTCCGAGAGAAAAGTTGTCTTTAGGTGTGACGATATTA 1174
QY	841 CTGTCTGTGACGGTGTCTCTCAACATGTTGGGAGACGATGCCAGACGCTCGACGCC 900
DB	1175 CTGTCTGTGACGGTGTCTCTCAACATGTTGGGAGACGATGCCAGACGCTCGACGCC 1234
QY	901 GTGCCCTTGTCTCGGCACCTACTTCAACTGCATCATGTTTCATGGTGGTTCCTCCCTCGTC 960
DB	1235 GTGCCCTTGTCTCGGCACCTACTTCAACTGCATCATGTTTCATGGTGGTTCCTCCCTCGTC 1294
QY	961 TCCACCATATGATCTCTCAACTACCAACCGGACGAGACACTCAAGAAATGATGAT 1020
DB	1295 TCCACCATATGATCTCTCAACTACCAACCGGACGAGACACTCAAGAAATGATGAT 1354
QY	1021 TGGATTGTTGGTGTCTCTTATTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
DB	1355 TGGATTGTTGGTGTCTCTTATTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1414
QY	1081 TCGGCGAGCGCGCGCGCGCGCTACTCTCGCGCGCGGAGCTCGAGCTGGCGGAG 1140
DB	1415 TCGGCGAGCGCGCGCGCGCGCTACTCTCGCGCGCGGAGCTCGAGCTGGCGGAG 1474
QY	1141 CGCTCTTCAAAGTCTCTTAGCGAAGCTGTCTCGACATCGATGAGACTTCGCGCACCCG 1200
DB	1475 CGCTCTTCAAAGTCTCTTAGCGAAGCTGTCTCGACATCGATGAGACTTCGCGCACCCG 1534
QY	1201 CAAGCGAGCAGCGCATGCTCGCATCTACAGGGGGGTGAGGAGAAATGGCGCGGG 1260
DB	1535 CAAGCGAGCAGCGCATGCTCGCATCTACAGGGGGGTGAGGAGAAATGGCGCGGG 1594
QY	1261 TTGGCGGCGCAGCTTGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB	1595 TTGGCGGCGCAGCTTGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1654
QY	1321 AGAGTCATCAGATCAGATCGCGCAAGAGACGAGATGCGGACATTTTCGCGGAGCTGG 1380
DB	1655 AGAGTCATCAGATCAGATCGCGCAAGAGACGAGATGCGGACATTTTCGCGGAGCTGG 1714
QY	1381 AAGTTCCGCGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB	1715 AAGTTCCGCGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1774
QY	1441 ATGCCACGCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1488
DB	1775 ATGCCACGCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1822

RESULT 4	AF321446	2023 bp	mRNA	linear	INV 29-APR-2002
LOCUS	AF321446				
DEFINITION	Drosophila melanogaster nicotinic acetylcholine receptor Dalphaf subunit variant type II (nAChRalpha-30D) mRNA, complete cds, alternatively spliced.				
ACCESSION	AF321446				
VERSION	AF321446.1	GI:20152846			
KEYWORDS	Drosophila melanogaster (fruit fly)				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 2023)				

ORIGIN

Query Match	41.1%;	Score 611.8;	DB 3;	Length 1665;
Best Local Similarity	64.4%;	Pred. No. 1.ee-113;		
Matches 948;	Conservative 0;	Mismatches 517;	Indels 6;	Gaps 2;
<hr/>				
Qy	21	CTCGCACTTGGCGGCCCCGCCGTGTCGTGTCGTGTCGTCTGTGCCTCTCTGGCCGAGGGG	80	
Dd	52	CCGCGTGCCAGCCTGCTGTGCTGTGTTGTCCTGTTGAATCTTCTGGCATATTAAAGA	111	
<hr/>				
Qy	81	GGCACGCTGCCGGGTACACGAGAAGCGGTACTGCACCACCTATTGGACCACTACAACGT	140	
Dd	112	AAGCTGTCAAGACCTCATGAANAAGCGCTGCTGAACCATCTGCTCCAACCTACATAAC	171	
<hr/>				
Qy	141	ACTGGAGAGCCCGTGTCTAACGAGAGCGACCCGCTGCGAGCTCTCTTCGGCCTCAACGCT	200	
Dd	172	GCTGGAGCGACCCGCTGGCCAATGAATCGAGGCCCTTGGAGGTTAAGTTTCGGACTGACGCT	231	
<hr/>				
Qy	201	CATGCGAGATCATCGAGCTGGACGAGAAGAACCAAGCTTTTTTAATAACAAAACATCTGGGCTAAA	260	
Dd	232	GCAGCAGATCATCGAGTGGACGAGAAGAATCAGCTTCTTCATAACGAATCTTTGGCTTTC	291	

292	GT	TGGAGTGGAA	CGACTACAATCTGCGCTGGAA	TGAACGGAATACGGCGGGTCAAGA	351
321	TT	TAAAGAGTGCCAC	CCACAGACTATGGAAC	ACAGACGTCTTATGTACAACAGCGCGGA	380
352	TC	TACGAATCAGCC	CAACAAGCTGTGGAAGCCGACGTGCTCATGTACAACAGCGCGGA	411	
381	CG	AAGGGTTCGACAG	CACATATCCACGAACGTGTGTGTCGGAACAACGGCTCGTGCT	440	
412	TG	AGGANTTCGATGG	CACGTATCACACCGCGTGTGGTCAACATGGCGGCAGTTGCT	471	
441	GT	ACGTGCGCCCGG	CCATCTTCAAGAGCACCTTCAAGATCGACATCACTGGTTCCCTTT	500	
472	GT	ACGTGCGCCCTGG	TATCTTCAAGAGCATGCAAGATGGACATCACTGGTTCCCATTT	531	
501	CG	ACGACCAACAGAT	CGAGATGAAGTTTCGCGAGCTGGACTTATGATGGTTATCAGTTGA	560	
532	TG	ATGACCAACATTC	GGAATGAAATTCGGTAGTTGGACTTACGATGGAAATCAGTTGA	591	
561	TC	TACAACCTACAGG	ATGAAGGGGGGAGATATAGCAGTTTGTCTACGAATGGGAATG	620	

592	TTTGGTTTTAGTTCCGAAGATGGAGGGATCTTTCCGATTTTCAATCAAAATGGCGAGTG	651
QY	GGAGTTAAATAGGAGTCCCCGGCAAGCGCAACGAGATCTACTACAATGTGTGTCGGAGCC	680
Db	GTACTTGGCTTGCCATGCCGGGAAAGAATAACGATAGTCTACGCTGCTGCCCAAGACC	711
QY	ATACATCGACATCAGTTTTCGGTGGTGTATCCGGAGGAAAACGCTCTACTACTTCTTCAA	740
Db	ATATGTCGATATCACCTTTTACTATATACAAATTCGTCCGGTACATTATATTTTTCAA	771
QY	TCGTATCGTCCCTCGTGCTCATCGCCTCCATGGCTCTATTGGGGTTTCACCTTGCCTCC	800
Db	TTTAAATGTGGCCATGTGTCTAATCTCATCGATGGCCCTACTTGGGCTTTCACATTCGCCCC	831
QY	AGACTCCGGAGAAAAGTTGTCTTTAGGTGTGACGATATTACTGTGCTTTCGGGTGTTCTCT	860
Db	GGATTCCGGCGAAGAACTGACGCTGGGCGTAATCTATCTACTATCATTTAAACAGTATTTCT	891
QY	CAACATGTGGCGAGACGATGCCAGCACTGTCCGACCGCTGCCCTTGTCTCGGACACCTA	920
Db	AAACCTTGTGCGCGAGTCCATCCGACAAACGTCGATGCTGTCTCTCTATATAGGCACTTA	951

Db 472 GTACGTGCCCTCGTATCTTCAAGAGCACATGCAAGATGACATCACGTGTTCCCAT 531
QY 501 CGACGACCAACGATCGAGATGAAGTTTGGAGCTGGACTATGATGTTATCATGTTGGA 560
Db 532 TGATGACCAACATTCGGAATGAATTCGGATGTTGGACTTACGATGGAATTCAGTTGGA 591
QY 561 TCTACAACTACAGGATGAAGGGCGGAGATATAGCAGTTTGTGACGAATGGCGAATG 620
Db 592 TTTGGTTTGGATTCGGAAGATGGAGGGAICTTTCCGATTTTCAACAAATGGCGAGTG 651
QY 621 GGAGTTAATAGAGTCCCGGCAAGCGCAACGAGATCTACTACAACTGTTGTCGGAGCC 680
Db 652 GTACTTGTTCGATCGCGGAAAGAAATACGATAGTCTACGCTGCTGCGCCAGAACC 711
QY 681 ATACATCCACATCAGTTTGGGTGGTATCGGAGGAAAGCTCTACTACTTCTTCAA 740
Db 712 ATATGTCATATCAGTTTACTATACAAATTCGTGCGCGTACATATATATTTTCAA 771
QY 741 TCTGATCGTCCCTCGTGTCTCATCGCTCCATGCTCTATTGGGGTTTCACTTGCCTCC 800
Db 772 TTTAATTTGCCAATGTGCTAATCTCATCGATGCGCTTACTGGCTTCACATTCGCGCC 831
QY 801 AGACTCCGAGAAAGTTGTCTTTAGTGTGACGATATTTACTGTGTGACGGTGTCT 860
Db 832 GGATTCGGCGAGAAACTGACGCTGGCGTAACTATCTACTATCATTAACAGTATTTCT 891
QY 861 CAACATGTGTGGGAGATGCCAGCGACGCTCGACGCGCTGCGCTTGGTCCGACCTA 920
Db 892 AAACCTTGTGCGCGAGTCCATCGACAAACGTCGATGCTGTCTTTATAGGACCTA 951
QY 921 CTTCAACTGCATCATGTTTCATGTTGCTTCTCCCTGCTCTCACATACATGATCCTCAA 980
Db 952 CTTCAATTTGCATCATGTTTCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTCAA 1011
QY 981 CTACCAACGCGGACGCGACACTCAGCAATGATGATGATGATGATGATGATGATGATGAT 1040
Db 1012 CTACCAACGCGGACGCGACACTCAGCAATGATGATGATGATGATGATGATGATGATGAT 1071
QY 1041 TTATTTGCTGCGTGGTGTGCGCATGCTCAGCGCGCGCTGCGGACGCGCGCGCC 1100
Db 1072 ACAATGGTGCCTGGATCTTTCGGAATGGTTCGACCGCGCTGCAAGATTACACGCAAAAC 1131
QY 1101 GCGCGCGCTA---CCTCGCGCGCGGAGCTGAGCTGCGGAGCGCTCTCCAAAGTCGCT 1157
Db 1132 AATACTATTAGCAATCGCATGAAGAGCTGGAGCTAAGGAGCGCTCTCCAATCCCT 1191
QY 1158 CTTAGCGAAGCTGCTGACATGATGACGACTTCGCGCAACCGCAAGCGGACGCGCGGA 1217
Db 1192 GCTGCGCAATGTCTTCGACATCGACGAGACTTTCGCGCACACAATATCTGGCTCCCAAAC 1251
QY 1218 ATGCTGCCGATACACGCGGGGTGAGGAGATGCGCGGGTGGCGGCGCAGATTG 1277
Db 1252 CGCAATGCTGCTGCGCGCAGCTTCGCTGCGCGCCACACGCTGGAGGAGCATCACACGGC 1311
QY 1278 CTTGCGGTGCG---ACTACGAGCTCTCCCTCATCTGAAGGAGATTAGAGTATCATCACA 1334
Db 1312 CATCGGCTGCAATCACAAGATCTTCATCTAATCTCAAGAAATTTGCAATTTATACGSC 1371
QY 1335 TCAGATGCGCAAGGACGAGATGCGGACATTCGCGCGGAGCTGCGGAGTTGCGCGCCAT 1394
Db 1372 GCGGATGCGCAAGCTGACGAGCGGAAATGATCGCGCGATTTGAAAGTTTGGCGGCAAT 1431
QY 1395 GGTCTGTGACAGCTGTGCTTATTATCTTTACCTGTTTCAAAATCATCGCCAGCCTAGC 1454
Db 1432 GGTGTGATAGATTTTGTAAATGTTTTCACGCTCTTCAGCATTTATGCAAGCTTAC 1491
QY 1455 GGTGTGCTGTCCGCGCACACATCATGCTG 1485
Db 1492 GGTGTGCTGTCCGCTCCGACATAATCGTG 1522

RESULT 7

AF321447 2023 bp mRNA linear INV 29-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6
DEFINITION subunit variant type III (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
AF321447
AF321447.1 GI:20152848
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2023)
AUTHORS Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
TITLE Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalphas, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
21969411
11973307
JOURNAL 2 (bases 1 to 2023)
MEDLINE Grauso, M. and Sattelle, D.B.
PUBMED Direct Submission
AUTHORS Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
TITLE Genetics Unit, University of Oxford, South Park Road, Oxford OX1
JOURNAL 3QX, UK
FEATURES Location/Qualifiers
source
1..2023
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1..2023
/gene="nAcRalpha-30D"
379..1863
/gene="nAcRalpha-30D"
/note="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3b and 8b"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit
variant type III"
/protein_id="AA013394.1"
/db_xref="GI:20152849"
/translation="MDSPLPASLSFLLIFLAIKESCQGPHEKRLNHLSTVNTL
ERPVAESEPLEVEFGLTQIIDVEKNQLLI TNLMSLEWYDNLNWNTEYGGVK
DLRITPNKLPDVLMTNSADEGDTYHTNIVVKGSGCLYVPPGIFKSTCKMDITW
PFDDQHCMEKFGSWYDGNLDLVLNSEDGDSDFITNGEWYLLAMPKQNTIYA
CCBPYVDITFTIOIRRTLYFFNLIVPCVLISSMALLGFTLPDPSGKLTIGVTIL
LSLVFLNVAETLPQSDAIPILGTYFNCIMFMVASSVLTIVVLYNHHRTADHEM
PFWKSVFLQWLPWLMRGRKTRITLLSNRKELEKRSKSLANVLDDDD
DREITSGSTGTAIGSSAFGRPTVEHTTAIGCNHKLHLILKELQFITARMRKADD
EAELIGWKFAAMVDFRLIVFTLTITATVTLLSAPHIVQ"
794
variation
/gene="nAcRalpha-30D"
/note="results in glycine to serine substitution; compared
to the sequence deposited in GenBank Accession Number
AE003626"
/replace="a"
846
variation
/gene="nAcRalpha-30D"
/note="results in asparagine to serine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"

ORIGIN

Query Match 41.1%; Score 611.8; DB 3; Length 2023;
Best Local Similarity 64.4%; Pred. No. 1.6e-113;
Matches 948; Conservative 0; Mismatches 517; Indels 6; Gaps 2;

[illegible]

QY	1101	GGCGGGGTA---CCTCGCGCGCGGACCTGGAGACTGGCGGAGCGCTCCTCCAAAGTCGCT	1115
DB	1467	AATACTATTAAAGCAATCGCATGAAGGAGCTGGAGCTAAAGGAGCGCTCCTCCAAATCCCT	1526
QY	1158	CCTAGCGAAACGTGCTCGACATCGATGACGACTTCGCGCCACCGCGCAAGCGAGCAGCGCA	1217
DB	1527	GCTGGCCAAATGCTCTGACATCGACGACGACTTCGCGCACACATATCTGGCTCCCAAC	1586
QY	1218	ATGCTGCCGATPACTACAGGGGGGTGAGGAGAAATGGCGCGGGTTCGCGCGCAGCTTG	1277
DB	1587	CGCCATTGGCTCGTCGGCGCAGCTTCGGTTCGCGCCCAACAAGGTGGAGGACATCACACGGC	1646
QY	1278	CTTCGGTGTGCG--ACTACGAGCTCTCCCTCATCTCTGAAGAGATTAGAGTCATCACAGA	1334
DB	1647	CATCGGCTGCAATCACAAAATCTTCACTTAATTTCTCAAGAAATTCGAATTTATTACGGC	1706
QY	1335	TCAGATGCGCAAGGACGACGAAGATCGGACATTTTCGCGCGACTGGGAAGTTTCGCGGCCAT	1394
DB	1707	GCGATGCGCAAGCTGACGACGAGCGGAATTGATCGCGAATTGGAAGTTTCGCGCAAT	1766
QY	1395	GGTCGTGGACAGACTGTGCGCTTATTATCTTTTACCCCTTTTCACAATCATCGCCAGCTAGC	1454
DB	1767	GGTTGTGGATAGATTGTTGTTTAAATGTTTTCACGCTCTTCACGATATTTCGAAGGTTAC	1826
QY	1455	CGTCTCTGTTCGCGGCCACACATCATGGTG	1485
DB	1827	GGTCTGCTCTCGCTCCGCACATAATCGTG	1857
RESULT 8			
AF321445			
LOCUS			
DEFINITION			
Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6			
subunit variant type I (nAcRalpha-30D) mRNA, complete cds,			
alternatively spliced.			
ACCESSION			
AF321445			
VERSION			
AF321445.1			
KEYWORDS			
SOURCE			
ORGANISM			
Drosophila melanogaster (fruit fly)			
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE			
AUTHORS			
TITLE			
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,			
Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a			
New and Highly Conserved Target of Adenosine Deaminase Acting on			
RNA-Mediated A-to-I Pre-mRNA Editing			
Genetics 160 (4), 1519-1533 (2002)			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional			
Genetics Unit, University of Oxford, South Park Road, Oxford OX1			
3QX, UK			
FEATURES			
source			
1. .2023			
/organism="Drosophila melanogaster"			
/mol_type="mRNA"			
/db_xref="taxon:7227"			
/chromosome="2"			
/map="30D1"			
/dev_stage="embryo"			
1. .2023			
/gene="nAcRalpha-30D"			
379, 1863			
/gene="nAcRalpha-30D"			
/notes="ion channel; neurotransmitter transmembrane			
receptor; alternatively spliced; contains exons 3a and 8b"			
/codon_start=1			
gene			
CDS			


```
Qy 1455 CGTCTGCTCTCCGCGCCACATCATCGTG 1485
||||| ||||| ||||| ||||| |||||
Db 1827 GGTGCTGCTCTCCGCGCCACATCATCGTG 1857

RESULT 9
AF272778 2907 bp mRNA linear INV 29-APR-2002
LOCUS
DEFINITION Drosophila melanogaster nicotinic acetylcholine receptor Dalphas
subunit (nAcRalpha-34E) mRNA, nAcRalpha-34E-A allele, complete cds.
ACCESSION AF272778
VERSION AF272778.1 GI:20152839
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 2907)
Grauso,M., Reenan,R.A., Culetto,E. and Sattelle,D.B.
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalphas, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
21969411
MEDLINE 11973307
PUBMED
REFERENCE 2 (bases 1 to 2907)
AUTHORS Grauso,M. and Sattelle,D.B.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2000) Human Anatomy and Genetics, MRC-FGU,
University of Oxford, South Park Road, Oxford OX1 3QX, UK
FEATURES
source
1..2907
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="2L; 34E4-5"
/dev_stage="embryo"
1..2907
/gene="nAcRalpha-34E"
/allele="A"
289..2712
/gene="nAcRalpha-34E"
/feature="ion channel; neurotransmitter transmembrane
receptor"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalphas
subunit"
/protein_id="AAMI3390.1"
/db_xref="GI:20152840"
/translation="MKNAQLKLTVDDEDLWLAVLAHCSNFSFSSSTRSTSSNORH
NMLSPKTAATAAGDRAATCOPTNIRLCARLRRKRPATRNEDTKKQQL
SMPPFKRSTDTYSTPAATSCPTATYMQCRASDNESFPI SHDRVSTATFANVLH
VLQVSLGLNQWQHVVQORSVLLFRFAASTIAFISYLGSAQLKNSSSSSSSSS
NNSQTQLNGLNKHSTFLFLIYLNLAKVLAHYEKRLLHLLDLPYTLERPVINES
DLQLSFGLTLMQIIDVDEKNQLLVNWLKLEWDMNLRWNTSDYGVKDLRPPHHR
IKPDLVLYNSADSGDTGYOTNVVVRNNGSLVVPPIFKSTCKIDITWPFDDORC
EMKFGSTYDGFOLDLOLQDETGGDISYVLNGEWELLGVPGKENELYNCCPEPYD
ITFALLIRRTLYEFNLIIPCVLIASWALLGFTLPDPSSEKLSLGVITLLSLVFLN
MVAETMATSDAVPLLGYFCIMFMVASSVVSTILLINVHNRNADTHEMSEIRIVF
LCWPLWLRSPKRLILEPPTPCSTDSERKHQILSDVELKERSKSLANLVDI
DDDRHNRMPWTPGGLPHNPAFYRTVYGGDDGSGIGFSTRMPDAVTHHTCIKSST
EYELGLILKEIRFTDQLRKDECDNDIANDWKFAAMVVDRLCLLIIFTMFAIIAVL
LSAPHIIVS"
373
/variation
/gene="nAcRalpha-34E"
/feature="compared to B allele"
/replace="a"
375
/variation

/gene="nAcRalpha-34E"
/feature="results in phenylalanine to isoleucine
substitution; compared to B allele"
/replace="c"
827
/variation
/gene="nAcRalpha-34E"
/feature="results in threonine to isoleucine substitution;
compared to B allele"
/replace="t"
1079..1080
/variation
/gene="nAcRalpha-34E"
/feature="results in lysine to arginine substitution;
compared to B allele"
/replace="gg"
1081..1086
/variation
/gene="nAcRalpha-34E"
/feature="results in deletion of asparagine and serine;
compared to B allele"
/replace="n"
1119..1120
/variation
/gene="nAcRalpha-34E"
/feature="results in serine, glycine, and asparagine
insertion; compared to B allele"
/replace="agcggcaac"
1121
/variation
/gene="nAcRalpha-34E"
/feature="results in asparagine to serine substitution;
compared to B allele"
/replace="g"

Query Match 39.2%; Score 583.8; DB 3; Length 2907;
Best Local Similarity 65.2%; Pred. No. 7.7e-108;
Matches 976; Conservative 0; Mismatches 422; Indels 99; Gaps 4;

Qy 91 GGGTACCACGAGAGCGCTACTGCACACCTATTGACCACTACAGTACTCGAGAGG 150
||||| ||||| ||||| ||||| |||||
Db 1213 GGATATCATGAAAGAGACTGTTACAGATCTTTGGATCCCTTATATACACTAGAACGT 1272

Qy 151 CCCGCTCAACGAGAGCGACCCGCTGCAGCTCTCTCGGCCCTCAGCTCATGCAGATC 210
||||| ||||| ||||| ||||| |||||
Db 1273 CCCGTTCTCAATGAATCGGACCCGTTACATTAAGCTTTGGTTTAATGCAATTT 1332

Qy 211 ATCGACGTGACGAGAGAACACAGCTTTTATATAACAACATCTGGCTAAACTAGATGG 270
||||| ||||| ||||| ||||| |||||
Db 1333 ATCGATGTGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAAACTGGAGTGG 1392

Qy 271 AATGATATGAATTTGAGTGGAAACACTTCAGATTTCCGCGGGGTCAAAGATTTAAGAGTG 330
||||| ||||| ||||| ||||| |||||
Db 1393 AACGACATGAATCTCCGCTGGAAACACTCCGACTATGCGGGAGTTAGGATCTCGGAATA 1452

Qy 331 CCACCCACAGACTATCGAAACCCAGACGTCCTTATGTATCAACAGCGCGGACGAGGGTTC 390
||||| ||||| ||||| ||||| |||||
Db 1453 CCGCGCATCGACTCTCGAAGCGGACGTCGTGATGTACAACAGTGCGGATGAGGGATTT 1512

Qy 391 GACAGACGCTATCCAAACGACGTTGGTGGGAGAACACCGCTGCTGTCTGTACGTGCCG 450
||||| ||||| ||||| ||||| |||||
Db 1513 GACGCGCACTACCAAGACGACGTTGGTGGTGGGAAACACCGCTGCTGTCTATACGTTCCG 1572

Qy 451 CCCGGCATCTTCAAGACACTCGCAAGATCGACATCACTGGTTCCTCCCTTCGACGACCAA 510
||||| ||||| ||||| ||||| |||||
Db 1573 CCGGGATCTTCAAGTCGACGTCGAAGATCGACATCACTGGTTCCTCCCTTCGATGACGAG 1632

Qy 511 CGATCGAGATGAAGTTTGGCAGCTTATGATGTTTATCAGTTTGGATCTCAACTA 570
||||| ||||| ||||| ||||| |||||
Db 1633 CCGTGCAGATGAAGTTTCGCGAGTTGGACCTACGACGATTCAGCTGGATTTACAATTA 1692

Qy 571 CAGATGAAGGGCGGAGATATAAGCAGTTTGTACGAATGSCGAATGGGAGTTAATA 630
||||| ||||| ||||| ||||| |||||
Db 1693 CAAGATGAAACTGGCGGTGATATCAGCAGTTTACGTCTCAACGCGAGTGGGAATCTG 1752

Qy 631 GGAGTCCCGGACGACGACGATCTACTACAACTGTTGTCCGGAGCCATACATCGAC 690
||||| ||||| ||||| ||||| |||||
```

Db 1753 GGTGTGCCGGCAACGTAACGAGATCTATTACAACTGCTGCCGGAACCCCTATTATAGAC 1812
QY ATCACTTTGGGTGGTGGTATCGGAGGAAGCGTCTACTACTTCTTCAATCTGATCGTG 750
Db 1813 ATCACTTTGGGTGGTGGTATCGGAGGAAGCGTCTACTACTTCTTCAATCTGATCATA 1872
QY CCCTGCTGCTCATCGCTCCATGCTCTATTGGGGTTCACCTTCCCTCCAGACTCCCGGA 810
Db 1873 CCTTGTGTAATGCTTCATGCTTCATGCTTCGATTCACCTTCGCGCGGATTCGGGT 1932
QY 811 GAAAGTTGCTTTAGGTGTGACGATATATCTGCTTGACGGTCTTCTCAACATGCTG 870
Db 1933 GAAAAATATCGTGGGTGTACCACTTCTGCTGACCGTGTCTTCTGAATATGCTT 1992
QY 871 GCGGAGAGATGCCAGCGAGCTGCGACGCGTGCCTTGTCTCGGACCTACTTCAACTGC 930
Db 1993 GCCGAGACAATGCCCGCTACTTCCGATCGGTGCCAATGCTGGGTACATATTTCAAATGC 2052
QY 931 ATCATGTTTCATGGTGGTTCCTCCGTCGCTCTCCACATGATCTCAACTTACCACAC 990
Db 2053 ATATGTTTATGGTAGCTTCATCCGTTGTGTCAACGATTTTAATATTAATCATCAT 2112
QY 991 CGGACGAGACACTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1050
Db 2113 CGAAATGCTGATACGACGAAATGTCGAAATGATGATGATGATGATGATGATGATGATG 2172
QY 1051 CGGTGGGTGCTGCGCATGTACGCGCCGG- - - - -CTCGCGGAGACGCGCC 1097
Db 2173 CATGATGATTTGCGAATGATGCTGCCGAGACGACGCTGATCTAGATTCCTCCGAC 2232
QY 1098 GCGGCGCGCGTACTCTCCGCG- - - - -CCGACCTGAGCTGGCG 1137
Db 2233 CCCTGTTCCGACACATCTCCGAGCGGAAGACACAGATCTCTCGACGTTGAGCTGAA 2292
QY 1138 GAGCGCTCTCAAGTCTCTTACGGAAGTCTGATGATGATGATGATGATGATGATGATG 1197
Db 2293 GAGCGCTCTCGAAATGCTGCTGCGCAACGCTACTAGACATGATGATGATGATGATGATG 2352
QY 1198 - - - - -CCGAGCGCAGCAGCGCAATGCTGCCGATACACAGCGGGGG 1241
Db 2353 AATGTGCGCCCATGACGCGCGGGAACATGCGCACACACCGCGCTTCTATCGCAG 2412
QY 1242 TGAGGAGAAATGCGCGGGTGGCGGCGCACAGTTGCTTCGGTG- - - - - 1285
Db 2413 GTTTATGACAAGGCGAGTGGCAGCATGCGGCAATGCGCAGCACCGGAATGCCGAT 2472
QY 1286 - - - - -TCGACTAGCAGCTCTCCCTCATTTCTG 1311
Db 2473 GCGGTACCCCATCATACGTGTCATCAATCAATCAATCAATCAATCAATCAATCAATCA 2532
QY 1312 AAGGAGATTAGATCATACAGATCAGATGCGCAAGGAGCAGAGATGCGGACATTTTCG 1371
Db 2533 AAGGAAATTCGCTTTTAACTGATCAGCTACGTAAGATGAGGATGCAATGACATTCG 2592
QY 1372 CGGACTGGAATTCGCGGCAATGCTGTCGACAGATGTCCTTTATTTATTTTACCCTG 1431
Db 2593 AATGATTGGAATTTGACGATGCTGCTGACAGCTGTCCTTATCATATTCACAATG 2652
QY 1432 TTCAAATCATGCGCAGCTAGCGTGTCTGTCGCGGCGACACATCATGTCGTG 1498
Db 2653 TTCGCAATATTAGCCCAATAGCTGTAATCATATCAGCACCACATATTATTGTCG 2709

RESULT 10
AF143847
LOCUS AF143847 3029 bp mRNA linear INV 27-MAY-1999
DEFINITION Heliothis virescens putative nicotinic acetylcholine receptor alpha
7-2 subunit mRNA, complete cds.
ACCESSION AF143847
VERSION AF143847.1 GI:4895006
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
1 (bases 1 to 3029)
Schulze, T., Oellers, N. and Adamczewski, M.
Putative alpha subunits of insect nicotinic acetylcholine receptors
more similar to vertebrate alpha 7 subunits and C. elegans Ce21
than to other insect nicotinic acetylcholine receptor alpha
subunits
Unpublished
2 (bases 1 to 3029)
Schulze, T., Oellers, N. and Adamczewski, M.
Direct Submission
Submitted (19-APR-1999) ZP-BTB, Bayer AG, Bldg. Q 18., Leverkusen
51368, Germany
Location/Qualifiers
1. 3029
/organism="Heliothis virescens"
/mol_type="mRNA"
/db_xref="taxon:7102"
95-1600
/note="hvnachra7-2"
/codon_start=1
/product="putative nicotinic acetylcholine receptor alpha
7-2 subunit"
/protein_id="AAD32698.1"
/db_xref="GI:4895007"
/translation="MAPMLAALALLPVSEQGPHEKRLNALLANNVTLERPVANE
SEPLEVRLGLTQIIDVDKQQLITNLSLEWDLNLRKNSDSEYGVKDLRLTPN
KLWPDVLMNSADEGFDGTQTVNVRSGSLDYPGPIKAWFPEDDQH
CDMKFGWTDGNQDLVLKDEAGLSDFTINGEWLIGMFKKNTITYACCEPYV
DVTFTIMRRLTYFPNLI VPCVLISSMALLGTLPPDSEKLTGLVTLISLTVFL
NLVAETLPOVSDAIPLLGTXENCIMEMVASSVLTVVVLYNHHRTADLHEMPOMIKSV
FLWMLPILNRSFGKXITRTKIMNTRMRELELKERSSKSLIANLVLDIDDPHGGP
PPNSTASTGNLGPSCIFRTDFRRSFVRPSTMEDVGGSLGSHRELHLILRELQPIA
RMKKADEAEELISDWKFAAMVVDVRFCLFVFTLTITIAVAVLLSAPHIVQ"
ORIGIN
Query Match 38.6%; Score 575; DB 3; Length 3029;
Best Local Similarity 64.6%; Pred. No. 4.6e-106;
Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;
QY 97 CACGAGAGCGGCTACTGCACCCCTATTGGACCCTACAACTACTGAGAGGCGCCGTC 156
Db 158 CACGAGAGAGACTCCTGAACGCGTGTCTGCGAACTACAAACCCCTGAGCGACCGGTG 217
QY 157 GTCAACGAGAGCGACCGCTGAGCTCTCTTGGCCTCAGCTCATCGTCATCATCATGAC 216
Db 218 GCCAACGAGAGCGAACCGCTAGAGGTGAGGTTCGGCTTGCCTGACGCAATCATTCAT 277
QY 217 GTGACGAGAGAAACGAGCTTTTAAACAAACATCTGGCTAAACTAGAGTGAATGAT 276
Db 278 GTGACGAGAGATCACTACTTATTAACCAATATATGGCTGTCTGTTGAGTGAATGAC 337
QY 277 ATGAACCTTGGGTGAAACACTTCAGATTTCGCGGGGTCAAAGATTTTAAGAGTGCACCC 336
Db 338 TACAACCTGAGGTGAAACGAGCGAGTATGCGGGGTCAAAGACCTCAGGATCAGCCC 397
QY 337 CACGACTATGAAACGAGAGCTCTTATGTACACGCGGAGCAAGAGGTTCACAGC 396
Db 398 AACAGTTGTGAGAGCGGAGCTCTTATGTATATAGTGTGAGAGGGTTTTGACGG 457
QY 397 ACCTATCCAAACGAGCTGGTGTGCGGAAACAGGCTGCTGTCTAGTGTGCGCGCCGCG 456
Db 458 ACCTACCAGACCAAGTGGTGTGAGAGCGCGGCGAGTTCCTGTAGTGTCCACCTGCG 517
QY 457 ATCTTCAAGAGACCTGCAAGATCGATCATCTGTTTCCCTTCGAGCAGCAACAGTGC 516
Db 518 ATATTCAAGAGACATGCAAGATGGACATCGCGTGTTCCTTCGAGCAGCAACACTGT 577
QY 517 GAGATGAGTTTGGCAGCTGACTTATGATGTTATGATGTTGATCTCAACTACAGGAT 576
Db 578 GATATGAAGTTCCGTTAGCTGGACATATGACGCAATCATGTTGGATCTGGTCTGCTAAAGAT 637

QY 577 GAAGGGGCGGAGATATACGAGTTTGTACGAAATGCGAATGCGAGTAAATAGGATC 636
Db 638 GAGCAGCGCGGCTCTATCGGACTTATACAAATGGGAGTGTATCTAATAGGAATG 697
QY 637 CCGGGAAGCGCAACGAGATCTACTACAACTGTGTCCGAGCCATACATCGACATCAG 696
Db 698 CCAGGCAAAAGAACACAAATACATACGAGTGTCCGAGCCCTACGAGGAGTCAAC 757
QY 697 TTTGCGGTGTGTATCCGAGGAAACGCTCTACTACTTCTTCAATCTGATCGTCCCTGC 756
Db 758 TTCACCAATCATGATGAGAGAGCAACCTGTACTACTTCTTCAACCTGATCGTCCCTGC 817
QY 757 GTGCTCATCGCTCTCATGCTCTATGCGGTTCACCTTGCCTCCAGATCCGAGGAAAG 816
Db 818 GTGCTGATCTCATGATGGCACTCTCGGCTTCACTGCGCAACAGATCCGAGAGAAA 877
QY 817 TTGTCCTTTAGGTGACGATATTAATCTGCTGTGACGTTTCCCTCAACATGTTGCGGAG 876
Db 878 CTACACTTGGAGTCACTATCTTCTATCGCTGACGTTTCCCTCAACCTGATGAGCGAG 937
QY 877 ACGATGCCAGCGATCGGAGCGCTGCCCTTGTCTCGGCACTTCACTTCAATCGATCATG 936
Db 938 ACCCTGCCACAGTCTCCGAGCTATCCCTGTTAGGAGCTACTTCAATGTCATCATG 997
QY 937 TTCATGTTGCTTCTCTCGTCTCCACATATCTTCACTTCACTTCACTTCACTTCACT 996
Db 998 TTCATGTTGCTTCTCTCGTCTCCACATATCTTCACTTCACTTCACTTCACTTCACT 1057
QY 997 GCAGACACTCAGAAATGAGTATGATGATGCTGCTGCTGCTTCTTCTTCTTCTTCTTCT 1056
Db 1058 GCTGATATACATGAAATGCCAGCTGATGATGATGATGATGATGATGATGATGATGAT 1117
QY 1057 GTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1113
Db 1118 ATACTGCGAATGTCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1177
QY 1114 CCGCGCGCGGACCTGAGCTGCGGAGGCTTCTTCAAGTCTTCTTCAAGGAGGAGGAG 1173
Db 1178 AGGATGAGGAGCTGGAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1237
QY 1174 GACATGATGACGACTTCCGCGCACCGCAAGCGCAGC-----AGCGCAATCTGCGG 1226
Db 1238 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1297
QY 1227 ATACTACAGGGGGTGGAGGAGATGCGGGGTGCGCGGCGACAGTGTCTTCG---- 1282
Db 1298 AATTGCGACCTGGGTGCTCAATATTCGCAACGATTTCCGCTCGGTGCTTCTGCTCGG 1357
QY 1283 -----GTGTCGACTACGAGCTTCCCTCATTT 1308
Db 1358 TCCACGATGGAAGACGTGGGCGCGGGTGGGTAGCCACCATCGCGAGCTGCACCTCATA 1417
QY 1309 CTGAAGGAGATTAGAGTATCAGATCAGATGCGCAAGGAGGAGGAGGAGGAGGAGGAGG 1368
Db 1418 CTGAGAGGCTGAGTATCATCGCGCAGAGTGAAGAGGCTGATGAGGAGGCGAGCTG 1477
QY 1369 TCGCGGAGCTGGAAGTTGCGCGCATGCTGCTGAGCAGAGTGTCTTATTTTATCTTTACC 1428
Db 1478 ATCAGGACTGGAAGTTGCTGCGATGCTTGTATAGGTTTTCCTGTTCTGCTGCTACA 1537
QY 1429 CTGTTCAATATCGCACGCTAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1485
Db 1538 CTTTTCATATCGGAGAGTASCTGTCTGTTATCGGACCGCATATCATCTGT 1594

RESULT 11
E58348
LOCUS E58348 3109 bp DNA linear PAT 18-JUN-2001
DEFINITION Nucleic acid encoding insect actyl choline receptor subunit.
ACCESSION E58348
VERSION E58348.1 GI:13019347
KEYWORDS JP 2000023680-A/3.

SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3109)
AUTHORS Martin, A., Nadja, E. and Thomas, S.
TITLE Nucleic acid encoding insect actyl choline receptor subunit.
JOURNAL Patent: JP 2000023680-A 3 25-JAN-2000;
BAYER AG
COMMENT OS Heliothis virescens
PN JP 2000023680-A/3
PD 25-JAN-2000
PF 26-APR-1999 JP 1999118159
PR 04-MAY-1998 DE 19819829.9
PI MARTIN ADAMUTSUEUSUKI, NADJA ERASU, THOMAS SCHULTE PC
C12N15/09, A01K67/033, C07K14/705, C07K16/28, C12N1/21, C12N5/10, PC
C12Q1/68.
PC G01N33/15, G01N33/50// (C12N1/21, C12R1:19), C12N15/00, C12N5/00 CC

Query Match 38.6%; Score 575; DB 6; Length 3109;
Best Local Similarity 64.6%; Pred. No. 4.6e-106;
Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;
QY 97 CACGAGAGCGGTACTTSCACACCTTATTCGACCACTTACCAACGATCTGGAGAGCCCGTTC 156
Db 158 CACGAGAGAGACTCTCTGAAACCGCTTCTGTCGCAACTTACCAACCTTGGAGCGACCGGTG 217
QY 157 GTCAACGAGAGGACCCGCTGCGAGCTCTCTTGGGCTCAGCTCATGCGAGATCATCGAC 216
Db 218 GCCAACGAGAGCGAACCGCTAGAGTTCAGGTTTCGCTTGCACCAATCAITTCAC 277
QY 217 GTGACGAGAGAACACAGCTTTTAAATACAAACATCTGGCTTAAACCTAGAGTGGAATGAT 276
Db 278 GTGACGAGAGAAATCAACTTATTAACCAATATATGGCTGTCTGTTGGAGTGAATGAC 337
QY 277 ATGAACCTTGAGTGGAACTTTCAGATTTTCGCGGGGTCAAAGATTTTAAAGAGTCCCAACC 336
Db 338 TACAACCTGAGTGGAAACGACGAGCTATGGCGGGTCAAGGACCTCAGGATCACGCC 397
QY 337 CACGACTATGGAACACGAGCTCTTATGTACAAAGCGGAGCGAGGTTTCGACAGC 396
Db 398 AACAGTTGTGGAAGCGCGAGCTCTTATGTATTAATAGTGTCTGACAGGGTTTTCGACGG 457
QY 397 ACGTATCAACGAGCTGTGTGCGGAACAAACGGCTCGTGTCTGTACGTGCCCGCCGCG 456
Db 458 ACCTACGAGCAACGTTGTGTGAGAGCGCGGAGTTCCTGTACGTGCCACCTTGGC 517
QY 457 ATCTTCAAGAGCACTTCAAGATGACATCACTTGGTTCCTTCCCTTCGACGACCAACGATGC 516
Db 518 ATATTCAAGAGCATGCAAGATGACATCGGTGGTTTCCCTTCGACGACCAACACTGT 577
QY 517 GAGTGAAGTTTGGCAGCTTATGATGTTTATCAGTTGATCTTCACTACACTACAGAT 576
Db 578 GATATGAAGTTCGGTAGCTGACATATGACGGCAATCAGTTGGATCTGGTCTAAAAGAT 637
QY 577 GAAGGGGCGGAGATATAAGCAGTTTGTACCAATCGCGAAATGGGAGTTTAAATAGGATC 636
Db 638 GAGCAGCGCGGATCTATCGGACTTCAATCAATGCGGAGTGTATCTTAAATAGGATG 697
QY 637 CCGCGCAAGCGCAACGAGATCTACTACAACCTGTTTCCGAGGAGCCATACATCGACATCAG 696
Db 698 CCAGGCAAAAGAACACAAATACATACGCTGTCTGCGCGAGCCCTACGTGGAGCTCAC 757

FEATURES
source
ORIGIN
FH Key Location/Qualifiers
FT CDS (95).. (1597).
Location/Qualifiers
1..3109
/organism="Heliothis virescens"
/mol_type="genomic DNA"
/db_xref="taxon:7102"

697 TTTCGGTGGTATCCGAGGAAACGCTCTACTACTTCTTCAATCTGATCGTCCCTGC 756
758 TTACCATCATGATAAGAGACGAACCTTGACTACTTCTTCAACCTGATCGTCCGTGC 817
757 GTGCTCATCGCTCATGCTCTATTTGGGTTTCACTTGCCTCCAGATCCCGGAGAAAG 816
818 GTGCTGATCTCATCGATGGCACTCTCGGCTTCACTGCCACCAAGATCCCGGAGAGAA 877
817 TTGCTTTTGGTGTGACGATATTACTGCTGTGAGGTTGCTTCAACATGTTGGCGGAG 876
878 CTCACACTTGGAGTCACTATTCTTCTATGCTGAGGTTGCTTCAACCTGATGCCGAG 937
877 ACAGTCCGAGCAGCTCGGACCGCTGCTTCTGCTCGGACCTACTTCAACTGATCATG 936
938 ACCCTGCCACAGTCTCCGAGCTATCCCTGTTTGGGAGCTACTTCAATTTGATCATG 997
937 TTTCATGGTGGCTTCTCCGTGCTCCACCATGATGCTTCAACTACACACCGGAC 996
998 TTTCATGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057
997 GCAGACACTCAGAAATGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
1058 GCTGATATACATGAATGCCAGTGGATAAATCAGTATTTCTTCAATGTTGCCATGG 1117
1057 GTGCTGCCATGTACAGCCCGGCTCGCG---ACGACGCGCGCGCGCGCGCTACCT 1113
1118 ATACTGCAATGTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1177
1114 CGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1173
1178 AGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1237
1174 GACATCGATGACGACTTCCGCGCACCGGCAAGCGGAGC-----AGCGCAATGTCGCG 1226
1238 GATATTGATGATGACTTCAGACACGCGCCTCGCGCTCTTACAGTACTGCTCGACCGG 1297
1227 ATACTAGGCGGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1282
1298 AATTGGGACCTGGGTGCTCAATATTCGCGACGAGATTTCCGTCGGTCTGCTCGTCCG 1357
1283 -----GTGTCGACTAGGAGTCTCCCTCAT 1308
1358 TCCAGATGGAAGAGCTGGCGGGGCTGGGTAGCCACCATCGGAGCTGCACCTCAT 1417
1309 CTGAAGGAGATTAGAGTATCATCAGATCAGATGCGCAAGGACGACGAAGATGCGGACAT 1368
1418 CTGAGAGAGCTGCATTCATCAGCGCCAGGATGAAGAGGCTGATGAGGAGCGGAGCTG 1477
1369 TCGCGGAGCTGGAAGTTCCGCGCATGCTGCTGAGAGAGTGTGCTTATTTATCTTACC 1428
1478 ATCAGCGAGCTGGAAGTTTCTGCGATGTTGTTGATAGGTTTGGCTGTTGCTGTTTACA 1537
1429 CTGTTTCAATATCATCGGACGCTAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1485
1538 CTTTTCATATCATCGGACAGTACTGCTCTGTTATCGGACCGCATATCATCTG 1594

RESULT 12
AX009614
LOCUS AX009614 3109 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent EP0962528.
ACCESSION AX009614
VERSION AX009614.1 GI:9996846
KEYWORDS Heliothis virescens (tobacco budworm)
SOURCE
ORGANISM Heliothis virescens
Eukaryota, Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota, Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1
AUTHORS Adamczewski, M.D., Schulte, T.D. and Oellers, N.D.
TITLE Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL Patent: EP 0962528-A 5 08-DEC-1999;

FEATURES
source 1..3109
Location/Qualifiers
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"
95..1600
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07501.1"
/db_xref="GI:9996847"
/db_xref="REMBL:CA07501"
/translation="MAPLAAALALILPVSEQGPHEKRLNLIANLYNLTLPVANE
SEPLEVRFGLTLOQIDVDEKNQLITNWLSEWLDYNLRWNSYEGVXDLRTTPN
KLWKPDLVMSNADGPDGTQTNVVRSGCLYVPGI PKSTCKMDIAEPPEDOH
CDMKPGWTDGNQLDLVLDEAGDLSDFITNGEWILGMPGKNITITYACCPBYV
DVPTIMIRRTLYIFNLIVPCLISSMALLGFTLPDPSGKFTLGLTLLSLVFL
NLVAETLPQVSDAIPLLGTGYFCIMFVASSVLTVVLYNHHRTADTHEPQWIKSV
FLOWPLILRMSRPGKIKTRKIMNTRMRELEKRSKSLANVLDDPRHGP
PPNSTASTNLGPGCSIFRTDFRSFVRPSTMEDVGGGLGSHHRLHILRELOPITA
RMKADBEAELISDWKFAAMVVDVFLVFTLTIIATVAVLLSAPHIIVQ"

ORIGIN
Query Match 38.6%; Score 575; DB 6; Length 3109;
Best Local Similarity 64.6%; Pred. No. 4.6e-106;
Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;
QY 97 CACGAGAGCGGCTACTCCACCCTATTGACCACTACACGCTACTCGAGAGCGCCCTC 156
Db 158 CACGAGAGAGACCTCTGAAACGCTTGTCTGCGAACTACACACCCCTGGAGCGCGTG 217
QY 157 GTCAACAGAGAGCGCGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 216
Db 218 GCCAAGAGAGAGCGCTAGAGGTGAGGTTCGGTTCAGCTTCGACCAATCATTCAC 277
QY 217 GTGACGAGAGAGACCGAGCTTTTAAACAAACATCTGGCTTAAACATAGATGGAATGAT 276
Db 278 GTGACGAGAGAGATCAACTACTTATAACCAATATATGGCTGTCTGCTGAGTGAATGAC 337
QY 277 ATGACCTTGTGCTGACACCTTCAGATTTCCGGCGGCTCAAGATTTAAGATGCCACCC 336
Db 338 TACAACCTGAGGTGGAACGACGAGGATATGGCGGGTCAAGGACCTCAGGATCACGCC 397
QY 337 CACAGACTATGAAACACGAGCTCTCTATGTACAAACAGCGGAGGAGGCTTCACAGC 396
Db 398 AACAGTTGTGAGAGCGGACGCTCTATGTATTAATAGTGTGACGAGGTTTTCACCGG 457
QY 397 ACCTATCCAGAGAGCTGGTGGGAAACAAACGCTGCTGTCTGTAGTCCGCGCCGCGC 456
Db 458 ACCTACAGAGACCAACGCTGGTGCAGAAAGCGCGGCGAGTTGCTGTAGTCCACCTGGC 517
QY 457 ATCTTTCAAGAGCCTCTCAAGATCGATCATCCTGTTCCCTTCGAGACCAACGATGC 516
Db 518 ATATTCAAGAGCAGATGCAAGATGACATCCGCTGTTCCCTTCGAGACCAACCTGT 577
QY 517 GAGATGAAGTTTGGCAGCTGACCTTATGATGTTTATCAGTTTGTATCTCAACTACAGAT 576
Db 578 GATATGAAGTTTCGCTAGCTGACATATGACGCAATCAGTTGGATCTGGTCTAAAGAT 637
QY 577 GAAGGGCGGAGATATAGCAGTTTGTACGAAATGGGAGTGGAGTTAATAGGAGTC 636
Db 638 GAGGACGCGCGGATCTATCGGACTTCATAACAAATGGGAGTGGTATCTAATAGGATG 697
QY 637 CCGGCAAGCGCAACGAGATCTACTCAACCTGTTGTCGGAGGCATATCATCGATCATCAG 696
Db 698 CCAGGCAAAAGAACACAAATACATACGCTGTCGCCGAGCCCTAGCTGAGAGTCAAC 757
QY 697 TTTCCGCTGGTGGATCCGAGGAGAAACGCTCTACTACTTCTTCAATCTGATCGTCCCTGC 756
Db 758 TTCACCATCATGATAAGAGACGAAACCTTGTACTACTTCTTCAACCTGATCGTCCGCTGC 817
QY 757 GTGCTCATCGCTCCATGGCTCTATTGGGGTTTCACTTGCCTCGAGCTCCGAGAAAG 816

Db 818 GTGCTGATCTCATCGATGGCACTCCTCGGCTTCACACTGCCACAGACTCCGGAGAAA 877
QY 817 TTGCTTTTAGGTGACGATATTAAGTCTGTGTTGACGGTGTCTCAACATGGTGGCGGAG 876
Db 878 CTCACACTTGGAGTCACTATTCTTATCGCTGACGGTGTCTCTCAACCTGGTAGCCGAG 937
QY 877 ACGATGCACGACGCTCGGACGGCTGCGCTTGTCTCGGACCTTCACTCAACTGCATCATG 936
Db 938 ACCCTGCCACAGGTCTCCGACGCTATCCCCCTGTTAGGAGTACTTCAATTGTCATCATG 997
QY 937 TTCATGGTGGCTTCTCCGCTGCTCCACCATACTGATCTCTCAACTACCACACCGGCAC 996
Db 998 TTCATGGTAGCGTCTGTGGTACTGACTGTGGTACTCAATTACCACCATCGAACA 1057
QY 997 GCAGACACTCAGAAATGAGTATGATGCTGTTGGCTGTTCTTATATGGCTGCCGTGG 1056
Db 1058 GCTGATATACATGAATGCCACAGTGGATAAATCAGTATTCTCTACAAATGGTGGCCATGG 1117
QY 1057 GTGCTGGCGATGTCACGCGCCCGCTCGCG---ACGACGCGCGCGCGCGCGGTACCT 1113
Db 1118 ATACTGCGAATGTCGAGCCAGGAAGAAGATCACAGGAAGACTATAAATGATGAACACG 1177
QY 1114 CCGCGCGGACCTGGAGCTGCGGAGCGTCTCTCAAGTCTCTTACGGAAGCGTCTC 1173
Db 1178 AGGATGAGGAGCTGGAACTGAAGGAGAGGTCTGCGAAGTCTTGTGCGGAATGTTCTA 1237
QY 1174 GACATCATGACGACTTCCGCCACCGCAGCGGAGC-----AGCGGAATGCTGCCG 1226
Db 1238 GATAATTGATGATGACTTCAGACACGCGCCCTCGGCTCTTAACAGTACTGCGCTCGACCGG 1297
QY 1227 ATACTACAGGGGGGTGAGGAGAAATGCGGGGTTCGGCGCGCACAGTTGCTTCG--- 1282
Db 1298 AATTGGGACTGGTGCTCAATATCCGACGAAATTCGTCGGTCTGTCGTCGCGCCG 1357
QY 1283 -----GTGTGACTACGAGCTCTCCCTCAT 1308
Db 1358 TCCACGATGAAGACGTGGCGGGCTGGTAGCCACCATCGGAGCTGCACCTATA 1417
QY 1309 CTGAAGGAGATTAGATCATCAGATCAGATCGGCAAGGACGACGAAAGTGGGACATT 1368
Db 1418 CTGAGAGAGCTGCGATTTCATCAGCGCCAGGATGAAGAAGCTGATGAGGAAGCGGAGCTG 1477
QY 1369 TCGCGCACTGGAAGTTGCGCGCATGCTGTGACACAGCTGTCCTTATATCTTACC 1428
Db 1478 ATACGCACTGGAAGTTGTCGCAUGTTGTTGATAGTTTTCCTGTCGTTGCCA 1537
QY 1429 CTGTTCAATATPCGCCAGCTAGCGTCTGCTGCCGCGCACACATCATGGT 1485
Db 1538 CTTTTCACATATCGCGACAGTAGCTGCTCTGTTATCGCACCGCATATCATCTG 1594

RESULT 13
DME554210 1683 bp mRNA linear INV 01-APR-2003
LOCUS Drosophila melanogaster mRNA for nicotinic acetylcholine receptor
DEFINITION subunit Dalp7 (nAcRalpha-18C gene).
ACCESSION AJ554210
VERSION AJ554210.1 GI:29466436
KEYWORDS nAcRalpha-18C gene; nicotinic acetylcholine receptor subunit Dalp7.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1683)
AUTHORS Millar,N.S.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) Millar N.S., Department of Pharmacology,
University College London, Gower Street, London, WC1E 6BT, UNITED KINGDOM
FEATURES source Location/Qualifiers
1. .1683

/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1. .1683
/gene="nAcRalpha-18C"
26. .1654
/gene="nAcRalpha-18C"
/codon_start=1
/product="nicotinic acetylcholine receptor subunit Dalp7"
/protein_id="CAD86936.1"
/db_xref="GI:29466437"
/translation="MSFPQPHSLPEATANGRMVYGLGLLIIMPACAGPHEKRLIH
ALDNNLSERPVNESDPLQLSGLTLMQIIDVEKNQLLITNIWLKLEWMDNLEW
NSSEFGVDRLRIPPHRLKPDVIMYNSADEGDFYATVAVVVRNNGSCILYVPPGIFK
STCKIDITWPFDDQCEKMFQSDGFDQLDQDEAGGDISSFITNGEWLLGVP
KRNELYNCPPPYDITPAILIRKLTLYFFNLIVPCVLIASMALLGTLPLPDGSE
KLSLGVTLILLSLTVFLNMVAETWPAISDAVPLLGKYENCIMFVASSVSTILLVNH
HRNPDHSESEWIRVIFLWPLCILMRQEGQVGYECPFPSSSSSSAGKQQION
VELKRSKSLANLVIDDDFRNHRCSATLPHOPTYRTMYRQDDGSGVPGFA
GPVVDGLRHEAISHITLTSAAEYELALIKELRWITTEQLKKEDETSIDTRDWKFAAMV
VDRLLIIFLFTIIATLAVLFSAPHFIVSGVRG"
26. .131
/gene="nAcRalpha-18C"
132. .1651
/gene="nAcRalpha-18C"
/product="nicotinic acetylcholine receptor subunit Dalp7"
ORIGIN
Query Match 38.2%; Score 568.4; DB 3; Length 1683;
Best Local Similarity 65.0%; Pred. No. 9,9e-105;
Matches 9/9; Conservative 0; Mismatches 416; Indels 111; Gaps 5;
QY 91 GGGTACCACAGAGCGGCTACTGCACCACTATTGGACCACTACACGTAAGGAGG 150
Db 131 GGACCCATGAGAGCGGCTACTCCAGCCCTTCTGGACAACCTACACACCTGGAGCGT 190
QY 151 CCCGTGCTCAACGAGACGACCCGCTGCACTCTCCCTTCGGCCTCAGCTCATGCGATC 210
Db 191 CCGGTGCTCAATGAATCCGATCCATGCACTCGAGCTTCGGAATAACACATCATGCAGATT 250
QY 211 ATCGAGTGCACGAGAACGACGCTTTTAATAACAACATCTGGCTAAACTAGAGTGG 270
Db 251 ATCGATGTGACGAAAAGAAATCAACTGCTTATAACGAATATTTGGCTCAAAATGGAAATGG 310
QY 271 AATGATATGAATTTGAGGTGGAAACACTTCAGATTTCGGCGGGCTCAAAGATTTAAAGAGTG 330
Db 311 AACGATATGAATCTTCATGGAAATTCGAGTGAGTTCGGTGTGTGCGGATCTCGCAATT 370
QY 331 CCACCCACAGACTATGGAAACGAGAGCTCTTATGTACACAGCGCGGACGAGGGTTC 390
Db 371 CCGCCACATCGCTATGGAAACCGGATGTACTGATGTACAAAGTGCAGCGAGGGCTTC 430
QY 391 GACAGACGCTATCCACGACGCTGGTGGTGGGAAACGACGCTGCTGTCTGTAGTCCCG 450
Db 431 GATGGAACTGACCGCCAAATGGTGGTTCGCAATAATGCGGAGCTGTCTGTACGTACCG 490
QY 451 CCCGGATCTTCAAGAGCACTGCAAGATCGACATCACTGCTTCCTCCCTTCGACGACCAA 510
Db 491 CCAGGTATATTTAAGTCAAGTGAAGTGAAGTGAAGTTCATTCGAGCATCAG 550
QY 511 CGATGCGAGATGAAGTTTGGCAGCTGGACTTATGATGGTTATCAGTTGATCTTACAACTA 570
Db 551 AGATGTGAAATGAAATTTGGTTCGTGGACCTACGATGGGTTTCAGTTGGACCTGCAGTTG 610
QY 571 CAGATGAAGGGGCGGAGATATAACGAGTTTGTACGATGGGAAATGGGAGTTAATA 630
Db 611 CAGGACGAAAGCTGGTGGCGACATTTCTAGCTTTATAACCAAGGCGAAATGGGACTTGTTA 670
QY 631 GGAGTCCCCGGCAAGCGCAACGAGATCTACTACAACTGTTGTCCGGAGGCCATACATCGAC 690
Db 671 GGTGTGCCCGGTAAACGAAATGAAATCTACTATATTTGCTGCCAGAACCTTATATTGAC 730


```
QY 1191 CCGCCACCCGACGAGCGCGCAATGCTGCCGATCTACTACAGGGGGGTGAGGAGAA 1250
Db |||||
QY 1647 TCGGCACACAATATCTGGCTCCCAACCGGCATTTGGCTCGTCGCCAGCTTCGGTCGGCC 1706
Db |||||
QY 1251 TGGCGGGGGTTGGCGGCGCACAGTTGCTTCCGCTGTCGACTAC---GAGCTCTCCCTCAT 1307
Db |||||
QY 1707 CACAACGGTGGAGAGCATCACACGGCCATCGGCTGCAATCAAAAGATCTTCATCTAAT 1766
Db |||||
QY 1308 TCTGAAGGAGATTAGATCATCAGATCAGATGCGCAAGGACGACGAGATCGGACAT 1367
Db |||||
QY 1767 TCTTAAAGAAATTCGAATTTATTACGGCGCGGATCGCAAGCTGACGACGAGCGAATT 1826
Db |||||
QY 1368 TTCGCGGACTGGAGTTCCGCCCATGGTCGTGGACAGACTGTGCCTTATTCTTTAC 1427
Db |||||
QY 1827 GATCGGCGATTGGAAGTTCGGGCAATGGTTGTGGATAGATTTTGTTTAATTGTTTCAC 1886
Db |||||
QY 1428 CCTGTTCACAATCATCGCCACGGTAGCCGTGCTGCTCGCGGCCACACATCATGGTG 1485
Db |||||
QY 1887 GCTCTTCACGATTATTGCAACGGTTACGGTGCTGCTCTCGCTCCGACATAATCGTG 1944
Db |||||
```

Search completed: May 8, 2004, 02:41:22
Job time : 5903.18 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 10:08:00 ; Search time 598.456 Seconds
(without alignments)
10562.710 Million cell updates/sec

Title: US-09-303-232-3_COPY_335_1822

Perfect score: 1488

Sequence: 1 atggggggggggcgccgcg.....cgccacacatcatgtgtcg 1488

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1488	100.0	3700	3	Aaz24476 H. viresc
2	575	38.6	3109	3	Aaz24477 H. viresc
3	512.8	34.5	2886	3	Aaz24475 D. melano
4	470	31.6	936	4	ABL13733 Drosophil
5	458.2	30.8	1540	4	ABL07231 Drosophil
6	405.6	27.3	1509	4	Rac90380 Wild-type
7	405.6	27.3	1876	2	Aat48239 Neuronal
8	405.6	27.2	1876	8	ADAI0864 Human neu
9	405.2	27.2	1876	2	AAV12197 Human neu
10	405.2	27.2	1876	6	ABSS4875 Human neu
11	405.2	27.2	1876	6	ABV73248 Human neu
12	404	27.2	1509	3	RAC58395 Human PRO
13	404	27.2	1509	4	AAC90385 Mutant hu
14	404	27.2	1509	4	AAC90386 Mutant hu
15	402.4	27.0	1509	4	Rac90387 Mutant hu
16	394	26.5	1590	2	AAV44687 V274T var
17	386.8	26.0	1964	6	ABZ11298 Human pol
18	327.4	22.0	2769	2	AAT59196 Neuronal
19	312.2	21.0	2101	2	AAT59197 Neuronal
20	309.2	20.8	2277	2	AAQ03199 Human neu
21	309.2	20.8	2277	2	AAQ03087 Alpha 2 s
22	309.2	20.8	2277	6	ABSS4870 Human neu
23	309.2	20.8	2277	6	ABV73243 Human neu

24	309.2	20.8	2664	2	AAT48235 Neuronal
25	309.2	20.8	2664	6	ABK92165 Prostate
26	309.2	20.8	2664	8	ADAI0854 Human neu
27	309.2	20.8	2664	9	ADC71170 Human 205
28	309.2	20.8	2666	9	ADC71168 Human 205
29	302.4	20.3	803	4	ABL07799 Drosophil
30	292	19.6	1908	2	AAT48236 Neuronal
31	292	19.6	1908	8	ADAI0856 Human neu
32	290	19.5	1416	4	AAC90382 Chimeric
33	289.8	19.5	1771	9	ADB78675 Human nic
34	289.8	19.5	1771	9	ADB78664 Human nic
35	288.2	19.4	1771	9	ADB78674 Human nic
36	288.2	19.4	1771	9	ADB78676 Human nic
37	286.6	19.3	1771	6	ADB26345 Human mut
38	286.6	19.3	1771	9	ADB78665 Human nic
39	286.6	19.3	1771	9	ADB78666 Human nic
40	286.6	19.3	1771	9	ADB78678 Human nic
41	286.2	19.2	1809	5	AAS91552 DNA encod
42	286.2	19.2	2082	2	AAT59527 Alpha4 su
43	285.2	19.2	1654	6	ABSS4871 Human neu
44	285.2	19.2	1654	6	ABV73244 Human neu
45	285.2	19.2	1757	2	AAV12200 Human neu

ALIGNMENTS

RESULT 1

AAZ24476

ID AAZ24476 standard; cDNA to mRNA; 3700 BP.

XX

AC AAZ24476;

XX

DT 17-FEB-2000 (first entry)

DE H. virescens acetyl-choline receptor DNA from clone Hva7-1.

XX

KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; AChR; ds.

XX

OS Heliothis virescens.

XX

PH Key Location/Qualifiers

FT CDS 335..1825

FT /*tag= a

FT /*product= "acetyl-choline receptor"

XX

PN DE19819829-A1.

XX

PD 11-NOV-1999.

XX

PF 04-MAY-1998; 98DE-01019829.

XX

PR 04-MAY-1998; 98DE-01019829.

XX

PA (FARB) BAYER AG.

XX

PI Adamczewski M, Oellers N, Schulte T;

XX

WPI; 2000-014207/02.

DR

P-PSDB; AAY50815.
New nucleic acid encoding a nicotinic acetylcholine receptor from insects, used to identify potential insecticides.

XX

PS Claim 1a; Page 14-17; 26pp; German.

XX

CC This invention describes a novel nucleic acid (NA) encoding a nicotinic acetyl-choline receptor (I) from insects which can be used as an insecticide. Inhibitors of (I) interfere with neurotransmission. (I) (also vectors containing it, its regulatory regions, and antibodies directed against (I)-encoded proteins) are used to screen for: (a) plant protection agents that alter conductance of AChR, potentially useful as

CC insecticides, or (b) genes which encode polypeptides that are involved in
 CC formation of functionally related AChR in insects. (1) are also used to
 CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (11). This sequence encodes an acetyl-choline
 XX receptor isolated from *Heliothis virescens*

CC	Sequence	3700 BP; 893 A; 953 C; 944 G; 910 T; 0 U; 0 Other;
CC	Query Match	100.0%; Score 1488; DB 3; Length 3700;
CC	Best Local Similarity	100.0%; Pred. No. 0;
CC	Matches 1488; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATGGGGCGGGGGCGCGCTCGCACTTGGCGGGCGCGGGCGCTGCTGCTGCTG 60
Db	335	ATGGGGCGGGGGCGCGCTCGCACTTGGCGGGCGCGGGCGCTGCTGCTGCTG 394
Qy	61	TGCTGCTGCTGGCGGAGGGGGCGAGCTGGGGTACCAACGAGAGCGGCTACTCGAC 120
Db	395	TGCTGCTGCTGGCGGAGGGGGCGAGCTGGGGTACCAACGAGAGCGGCTACTCGAC 454
Qy	121	CTATTGACCACTAACAGTACTGGAGAGCGCGCTGCTCAACGAGAGCGCGCTCGAG 180
Db	455	CTATTGACCACTAACAGTACTGGAGAGCGCGCTGCTCAACGAGAGCGCGCTCGAG 514
Qy	181	CTCTCTTGGCGCTCAGCTCATCGAGATCATCGACGTGGACGAGAGAAACAGCTTTTA 240
Db	515	CTCTCTTGGCGCTCAGCTCATCGAGATCATCGACGTGGACGAGAGAAACAGCTTTTA 574
Qy	241	ATAACAACTCTGGCTAAACTAGATGAGTGAATGATATGAATTTGAGGTGGAACACTTCA 300
Db	575	ATAACAACTCTGGCTAAACTAGATGAGTGAATGATATGAATTTGAGGTGGAACACTTCA 634
Qy	301	GATTTGGCGGGGTCAAGATTTAAGAGTGCACCCACAGACTATGGAACACAGAGCTC 360
Db	635	GATTTGGCGGGGTCAAGATTTAAGAGTGCACCCACAGACTATGGAACACAGAGCTC 694
Qy	361	CTTATGTAACAGCGCGGAGAAAGGTTGACAGACAGTATCAACGAAAGCTGGTGGTG 420
Db	695	CTTATGTAACAGCGCGGAGAAAGGTTGACAGACAGTATCAACGAAAGCTGGTGGTG 754
Qy	421	CGGACAAACGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db	755	CGGACAAACGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 814
Qy	481	GACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db	815	GACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 874
Qy	541	TATGATGGTTATCAGTTGGATCTACAACTACAGGATGAAGGGCGGAGATATAAGCAGT 600
Db	875	TATGATGGTTATCAGTTGGATCTACAACTACAGGATGAAGGGCGGAGATATAAGCAGT 934
Qy	601	TTTGTACGAATGGCAATGGAGTTAATAGAGTCCCGGCAAGCGACGAGATCTAC 660
Db	935	TTTGTACGAATGGCAATGGAGTTAATAGAGTCCCGGCAAGCGACGAGATCTAC 994
Qy	661	TACAACTGTTTCCCGGAGCCATACATCGACATCACGTTTGGGTGGTGTATCCCGAGAGAA 720
Db	995	TACAACTGTTTCCCGGAGCCATACATCGACATCACGTTTGGGTGGTGTATCCCGAGAGAA 1054
Qy	721	ACGCTCTACTACTTCTTCAATCTGATCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTA 780
Db	1055	ACGCTCTACTACTTCTTCAATCTGATCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTA 1114
Qy	781	TTGGGGTTTCACTTGGCTCCAGACTCCGAGAGAAAGTTGCTTTAGGTGTGACGATATTA 840
Db	1115	TTGGGGTTTCACTTGGCTCCAGACTCCGAGAGAAAGTTGCTTTAGGTGTGACGATATTA 1174
Qy	841	CTGTGCTTGGAGGTTTCTTCAACATATGGTGGGAGACGATGCGCAGCGAGCTCGGAGGCC 900
Db	1175	CTGTGCTTGGAGGTTTCTTCAACATATGGTGGGAGACGATGCGCAGCGAGCTCGGAGGCC 1234
Qy	901	GTGGCCTTGTGCGGACCTACTTCACTGATCATGTTCATGTTGCTGCTGCTGCTGCTGCTG 960

Db	1235	GTGCCCCCTTGTCTGCGCACCTACTTCAACTGATCATGTTCTGTTGGCTTCTCTCGTCTGTC 1294
Qy	961	TCCACCACTACTGATCTCTCACTACCAACCGGCGACGACACTACGAAATGAGTGAT 1020
Db	1295	TCCACCACTACTGATCTCTCACTACCAACCGGCGACGACACTACGAAATGAGTGAT 1354
Qy	1021	TGGATTCGTTGGCGTTCCTTTATTGGCTGCCGTTGGTGTGCGCATGTACGCGCCCGGC 1080
Db	1355	TGGATTCGTTGGCGTTCCTTTATTGGCTGCCGTTGGTGTGCGCATGTACGCGCCCGGC 1414
Qy	1081	TCGCGCAGCAGCGCGCGCGCGCTTACCTCCGCGCGCGGACCTTGGAGCTCGCGGAG 1140
Db	1415	TCGCGCAGCAGCGCGCGCGCGCTTACCTCCGCGCGCGGACCTTGGAGCTCGCGGAG 1474
Qy	1141	CGCTCTCTCAAGTCTCTTCTAGCGAAGCTGTGATGACATGCGATGACGACTTCCGCGCACCG 1200
Db	1475	CGCTCTCTCAAGTCTCTTCTAGCGAAGCTGTGATGACATGCGATGACGACTTCCGCGCACCG 1534
Qy	1201	CAAGCGCAGCAGCGCGCAATGCTGCCGATCTACAGGGGGGTGAGGAGAAATGGCGCGGG 1260
Db	1535	CAAGCGCAGCAGCGCGCAATGCTGCCGATCTACAGGGGGGTGAGGAGAAATGGCGCGGG 1594
Qy	1261	TTGCGCGCGCACAGTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db	1595	TTGCGCGCGCACAGTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1654
Qy	1321	AGAGTATCATCAGATCAGATGCGCAAGGAGCAGCAAGATGCGGAGATTTGCGCGGACTGG 1380
Db	1655	AGAGTATCATCAGATCAGATGCGCAAGGAGCAGCAAGATGCGGAGATTTGCGCGGACTGG 1714
Qy	1381	AAGTTCCGCCCGCATGGTGGCGAGACTGTGCGCTTATTTACCTGTTTCCCAATC 1440
Db	1715	AAGTTCCGCCCGCATGGTGGCGAGACTGTGCGCTTATTTACCTGTTTCCCAATC 1440
Qy	1441	ATGCCACGCTAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1488
Db	1775	ATGCCACGCTAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1822

RESULT 2

AAZ24477
 ID AAZ24477 standard; cDNA to mRNA; 3109 BP.

XX AC AAZ24477;

XX DT 17-FEB-2000 (first entry)

XX DE H. virescens acetyl-choline receptor DNA from clone Hva7-2.

XX KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 KW neurotransmission; plant protection agent; conductance; AChR; ds.

XX OS *Heliothis virescens*.

XX PN DE19819829-A1.

XX PD 11-NOV-1999.

XX PF 04-MAY-1998; 98DE-01019829.

XX PR 04-MAY-1998; 98DE-01019829.

XX PA (FARB) BAYER AG.

XX PI Adamczewski M, Oellers N, Schulte T;

XX DR WPI; 2000-014207/02.

XX DR P-PSDB; AAY50816.

XX PT New nucleic acid encoding a nicotinic acetylcholine receptor from
 XX insects, used to identify potential insecticides.

xx Adamczewski M, Oellers N, Schulte T;
xx WPI; 2000-014207/02.
xx P-PSDB; AAY50814.
xx New nucleic acid encoding a nicotinic acetylcholine receptor from
xx insects, used to identify potential insecticides.
xx Claim 1a; Page 8-12; 26pp; German.
xx This invention describes a novel nucleic acid (NA) encoding a nicotinic
xx acetyl-choline receptor (I) from insects which can be used as an
xx insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
xx (also vectors containing it, its regulatory regions, and antibodies
xx directed against (I)-encoded proteins) are used to screen for: (a) plant
xx protection agents that alter conductance of AChR, potentially useful as
xx insecticides, or (b) genes which encode polypeptides that are involved in
xx formation of functionally related AChR in insects. (I) are also used to
xx isolate and characterize the specified regulatory regions and for
xx recombinant production of (II). This sequence encodes an acetyl-choline
xx receptor isolated from *Drosophila melanogaster*
xx SQ Sequence 2886 BP; 942 A; 687 C; 640 G; 617 T; 0 U; 0 Other;
Query Match 34.5%; Score 512.8; DB 3; Length 2886;
Best Local Similarity 62.5%; Pred. No. 5.4e-115;
Matches 878; Conservative 0; Mismatches 502; Indels 24; Gaps 4;
Qy 91 GGGTACACGAGAGCGCTACTCCACACCTATTGGACCACTACAACGTAAGGAGG 150
Db 1296 GGATATCATGAAAGAGACTGTACAGATCTTTTGGATCCTTATAATACATAGAAGT 1355
Qy 151 CCGCTCTGACGAGAGCGCCCTGAGCTCTCTTCGCTCAGCTCAGCTCAGTATC 210
Db 1356 CCGCTCTCAATGAATCGGACCGCTTACAATTAAGCTTGGTTTAACTTTAATGCAATT 1415
Qy 211 ATCGACGTGGACGAGAGAACCGACTTTTAAATACAAACATCTGGCTAAACAGAGTGG 270
Db 1416 ATCGATGGAGAGAGAAATCANTTCTAGTCACTAATGTGTGTTAAACAGTGGAGTGG 1475
Qy 271 AATGATATGAATTGAGGTGAAACATTCAGATTTCCGCGGGGTCAAGATTTAAGAGTG 330
Db 1476 AACACATGAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGAATA 1535
Qy 331 CCACCCACAGACTATGAAACAGAGCTCTTATGTACACAGCGCGGAGAGGCTC 390
Db 1536 CCGCGCATCGATCTGGAACCGGAGCTGTGTATGTACACAGTGGGATGAGGATTT 1595
Qy 391 GACAGCAGTATCCAAAGAGCTGGTGGTGGGAAACACCGCTCGTGTGTAGTCCG 450
Db 1596 GACGGCACCTACGACAGCGAGTGGTGGTGGGAAACACCGCTCGTGTATACGTTCCG 1655
Qy 451 CCGGATCTTCAAGAGACCTGCAAGATGACATACCTGTTCCCTGTCGACGACAA 510
Db 1656 CCGGGATCTTCAAGTCCAGCTGCAAGATGACATACCTGTTCCCTGTCGATGACAG 1715
Qy 511 CGATGCGAGATGAATTTGCGAGCTGACATATGATGTTATGATGTTGATCTACACTA 570
Db 1716 CCGTGCAGATGAATTTGCGAGTTGGACCTACACGAGTTCCAGCTGGATTTACAATA 1775
Qy 571 CAGATGAAGGGGCGGAGATATAAGCAGATTTTGTACGAATGCGGAATGGAGTTAATA 630
Db 1776 CAAGATGAATTTGCGGTGATATCAGCAGTTACGTCTCAACGCGAGTGGAACTACTG 1835
Qy 631 GGAGTCCCGCAGACCGAGAGATCTACTAAGCTGTTGTCGGGAGCCATACATCGAC 690
Db 1836 GGTGTCCCGCAACCGTAAGAGATCTATTTACAACCTGCTGCCCGAACCCCTATATAGC 1895
Qy 691 ATCAGTCTTGGGTGGTATCCGAGGAAACCGCTCTACTACTTCTCAATCTGATCGTG 750
Db 1896 ATCACTTTCGCATCATCCGCGGACGACACTGTACTTCTTCAACCTGATCATA 1955

Qy 751 CCCTGCTGTCTATCGCTCCATGGCTCTATTGGGGTTTCACTTGGCTTCCAGATCCGGA 810
Db 1956 CTTTGTGTACTGATGTGCTCCATGGCTTCTCGATTCGATTCACCTTGGCCGAGATTCGGGT 2015
Qy 811 GAAAAGTTGTTTATGAGTGTGACGATATTACTGTCTGTGAGCGTGTCTCTCAACATGTTG 870
Db 2016 GAAAATTTATCGTGGGTGTTTACCATCTTGTCTGTGAGCGGTGTTCTGAATATGTTT 2075
Qy 871 GCGGAGACGATCCGAGCGAGCTGGAGCGCGTGGCTTGTCTCGGACCTACTTCAATGTC 930
Db 2076 GCGGAGACATGCGCGCTACTTCCGATGCGGTGCCATTG-----TGGATACGC 2123
Qy 931 ATCATGTTCTATGTTGGCTTCTCCGTCGTCTCCACCATACTGATCTCACTCACTACACAC 990
Db 2124 ATCGTGTGTTG--TGTGGCTGCCATGGATATTGCGAATGAGTGGCCAGGACGCGC 2181
Qy 991 CGGCACGACGACACTCAAGAAATGAGTGTGATTTGGATTCTGTGCGTGTCTTTTATGCGTG 1050
Db 2182 TGATCTCTAGAGTTCCCGACCAACCGCTGTTTCGGACACAT-----CCTCCGAGCGGAAGCAC 2237
Qy 1051 CCGTGGGTGCTGCGCATGTACGCGCGCTCGCGGAGCGCGCGCGCGCGCGCGGTA 1110
Db 2238 CAGATACTCTCCGACGTTGAGCTGAAAGAGCGCTCGTCCGAAATCGCTGCTGGCCACAGTA 2297
Qy 1111 CCTCCGCGCGGACCTGGAGCTGCGGAGCGCTCTCCAAAGTCGCTCCTAGCGAACCTG 1170
Db 2298 CTAGACATCGATGATGACTTCCGGCACAATTTGTGCCCATGACGCGCGGCGGAACACTG 2357
Qy 1171 CTGACATCGATGACGACTTCCGCCACCCCGAAGCGAGCGAGCGCAATGCTGCCGATAC 1230
Db 2358 CCACACAACCGGGCTTTCTATCGCACGGTTTATGGAACAAGCGGACGATGGCAGCATTTGG 2417
Qy 1231 TACAGGGGGGTGAGGAGATGGCGGGGTTCGGCGGCGCACAGTTCCTTC-----GGT 1284
Db 2418 CCAATTGGCAGACCCCGAATGCGGATGCGGTCAACCCATCATAGTGCATCAAAATCATCA 2477
Qy 1285 GTGACTATCAGAGCTCTCCCTCAITTTCTGAAGAGATTAGAGTCAATCAGATCAGATCGC 1344
Db 2478 ACTGAATATGAATTAGGTTTAACTTTAAAGGAAATTCGCTTTATACTGATCAGTACGT 2537
Qy 1345 AAGGACGACGAAGATGCGGACATTTCCGCGCATGTTGGAAGTTCGCGCCATGTTGCTGGAC 1404
Db 2538 AAAGATCAGAGTGCATGACATTTGCCAATGATTTGGAATTTGAGCTATGTTGCTGTTGAC 2597
Qy 1405 AGACTGTCCTTATTATCTTTTACCTGTTCCAAATCATCGCAGCTAGCGCTGCTGCTG 1464
Db 2598 AGACTGTCCTTATTATCATATTACAAATGTTTCCAAATATTAGCCAAATAGCTGTACTA 2657
Qy 1465 TCCGCGCCACACATCATGTTGTCG 1488
Db 2658 TCGGACCAACATATTATTGTCG 2681
RESULT 4
ABL13733
ID ABL13733 standard; cDNA; 936 BP.
XX ABL13733;
AC ABL13733;
DT 26-MAR-2002 (first entry)
XX *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 35681.
DE *Drosophila*; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX *Drosophila melanogaster*.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
PF

Query Match	27.3%; Score 405.6; DB 4; Length 1509;
Best Local Similarity	56.5%; Pred. No. 6.7e-89;
Matches	840; Conservative 0; Mismatches 604; Indels 42; Gaps 3;
QY	35 CGCCCGGGCCTGCTGCTGCTGCTGCTGCGGAGGGGGGACGCTGCGGGT 94
DB	11 CGCCGGAGGCGTGTGCTGGCGCTGCCGCGTCTCTGCAAGTCTCCCTGCAAGGG 70
QY	95 ACCACGAGAAGCGCTACTGCACACACTATTGGACCACATAACAGTACTGGAGAGCCCG 154
DB	71 AGTTCCAGAGGAAGCTTACAAGAGAGTGTCAAGAACTACATCCCTTGGAGAGCCCG 130
QY	155 TCGTCAACGAGAGGACCGCGCTGAGCTCTCTTGGCGCTCAAGCTCATGAGATCATCG 214
DB	131 TGGCCAAATGATCGCAACCACTACCGCTCTACTTCTCCCTGAGCCTCTGAGATCATGG 190
QY	215 ACCTGACGAGAGAACACCACTTTTAATAACAACAATCTGGCTTAAACTAGAGTGAATG 274
DB	191 ACCTGGATGAGAGAACCAAGTTTAAACCAACAATTTGGCTTGCAATGTTCTTGGACAG 250
QY	275 ATATGAACCTTGAGGTGGAACTTTCAGATTTCGCGGGGTTCAAGAATTTAAGAGTGCCAC 334
DB	251 ATCACTATTACAGTGGAAATGTGCAAAATATCCAGGGGTGAAGACTGTGTGTTCCCGAG 310
QY	335 CCCACAGACTATGGAACACAGAGCTCTTAATGATCAACACGCGGAGAGAGGTTTCGACA 394
DB	311 ATGGCCAGATTTCGAAACACAGACATTTCTCTATTAACAGTGTGTGATGAGCGCTTTGACG 370
QY	395 GCAGGTATCAACGAGAGTGGTGGTGGGAACAACGCTCGTCTGATAGTGGCGGCCG 454
DB	371 CCACATTCACACTAAGCTGTGGTGAATTCCTTGGGCAATGCCAGTACTCGCTCCAG 430
QY	455 GCATCTTCAAGAGCACCTGCAAGATGCACATCACCTGGTTCCTCTTCGACGACCAACGAT 514
DB	431 GCATTTCAAGAGTTCCTGTCTACATCGATGTCAGCTGTTTCCCTTTGATGTGCGACT 490
QY	515 GCGAGATGAAGTTTGGCAGCTGGACTTTATGATGGTTATCAAGTTGGATCTACAACATCAGG 574
DB	491 GCAACTGAAGTTTGGTCTCTGTCTTACGAGGCTGCTCTTGGATCTGCAGATGCAGG 550
QY	575 ATGAAGGGGGGAGATATAAGCAAGTTTCTCAGAAATGGGAATGGAGTTTAATAGGAG 634
DB	551 A-----GGCAGATATCAGTGGCTATATCCCAATGGAAATGGACCTTAGTGGAA 601
QY	635 TCCCGGCAAGCGCAACGAGATCTACTACAATCTTGTTCGGAGCCATACATCGAATCA 694
DB	602 TCCCGGCAAGAGAGTGAAGGTTCTATGAGTCTCAAGAGGCCCTACCCGATGTCA 661
QY	695 CGTTTGGGTGGTATCCGAGGAAAACGCTCTACTACTTCTTCAATCTGATCGTCCCT 754
DB	662 CTTTCAAGTGAACCATGCGCGGAGGACGCTCTACTATGGCTCAACCTGCTGATCCCT 721
QY	755 GCGTCTCATCGCTCCATGGCTTATTTGGGTTACCTTGGCTCAGACTCCGAGAAA 814
DB	722 GTGTGCTCATCTCGCCCTCGCCCTGTGTGTCTTCTCTCTGAGATTCGCGGGAGA 781
QY	815 AGTTGTCTTTAGGTGTGAGATATTAAGTGTGTTGACGGTGTCTCTCAACATGGTGGCG 874
DB	782 AGATTTCCCTGGGATTAACAGCTTATCTCTCTTACCGTCTTATGCTGCTGTGGCTG 841
QY	875 AGAGATGCCAGGACGTCGAGAGCCGTCGCTTGTCTGGCACTCTCAACTGCAATCA 934
DB	842 AGATCATGCCGCAACATCCGATTCCGATACCATTGATAGCCAGTACTTGGCAGACCA 901
QY	935 TGTTCATGTGGCTTCTCGTCTCTCACCATACTGATTCCTCAACTACCAACCGCG 994
DB	902 TGATCATCTGTGGGCTCTCGGTGTGTGAGCGGTGATCGTGTGAGTACCAACCAACG 961
QY	995 ACGCAGACACTCAGAAATGATGATTGGATTCTGTGCTTCTCTTTATTTGCTGCGCT 1054
DB	962 ACCCGCAGCGGGGCAAGATGCCAAGTGGACAGAGTCACTCTTCTGAATGTGTGCGCT 1021
QY	1055 GGGTGTCTCGCATGTCAACGGCCCGCTCGGCGAGCGAGCGCGCGCGCGCTAGCTC 1114

[illegible]

QY	DB	974	TGATCATCTGGGGCCTCTCGGTGGTGGTACGGGTGATCGTGTGCAGTACCAACCAACG	103
QY	995	ACGCAGACACTCACAAATGAGTATTGGATTTCGTTGCGGTTCCTTTATTTGGCTGCGGT	1054	
DB	1034	ACCCGACGGGGCAAGTCCCAAGTGGACAGATCATCTTCTGAACTGGTGGCGGT	1093	
QY	1055	GGGTGTGGCGATGTCTACGCGCCGGCTTCGCGCAGCAGCCGCGCGGGCGCGGTACCTC	1114	
DB	1094	GGTTCTCTCGGAATGAAGAGCCCGGGGAGACAAAGGTGCGCCCGGCTGCCAGCACAAGC	1153	
QY	1115	CGCG-----CGGACCTGAGCTGCGCGAGCGCTCTCTCAAGTCGCTCCTAGCGA	1165	
DB	1154	AGCGCGCTGCAGCCTGCGCCAGTGTGGAGATGAGCGCGGTGGCGCCGCCCGCCAGCA	1213	
QY	1166	ACGTGTCTGATCGATGACGACTTCGCGCACCCGCAAGCGCAGCGCAGCCGAATGCTGCC	1225	
DB	1214	ACGGGAACCTGCTGTATCATCGGCTTCGCGGCTTGGACGGCGTGCACTGTGTCCCGACCC	1273	
QY	1226	GATACTACAGGGGGGTGAGAGAAATGGCGGGGTGGCGGCGCAGCTTGTTCGGTG	1285	
DB	1274	CGGACTCTGGGTAGTGTGTGGCGCATGTCCTCCCGCCAGCAGATGACACCTTC	1333	
QY	1286	TCGACTACGAGCTCTCCCTC-----ATTCTGAAGGAGATTA	1321	
DB	1334	TGCAGCGCGGAACCCCGAGGGGACCCGGACTTGGCCAAGATCTTGGAGGAGTCC	1393	
QY	1322	GAGTCATCAGATCAGATCGCATGCGCAAGCAGCAAGATGCGGACATTCGCGCGACTGGA	1381	
DB	1394	GCTACATGTCGAATCGTTCCGCTGCCAGGACGAAAGCAGGCGGTCTGCAGCGAGTGA	1453	
QY	1382	AGTTTCGCGGCATGTGCTGCGCAGACTGTCCTTATTCTTTACCTGTTCCACATCA	1441	
DB	1454	AGTTTCGCGGCTGTGTGTGACCGCTGTGCTCATGGCTTCTCGTCTTCCATCA	1513	
QY	1442	TCGCCAGCTAGCGGTGCTGCTCTCCGCGCCACATCATGTGTCT	1487	
DB	1514	TCGTGACCATCGGATCTCTGATCGGCTCCCAACTTCGTGGAGGC	1559	
RESULT 8	ADA10864	ADA10864 standard; DNA; 1876 BP.		
XX	AC	ADA10864;		
XX	AC	06-NOV-2003 (first entry)		
DT	DT	Human neuronal nicotinic acetylcholine receptor alpha 7 subunit DNA.		
DE	DE	ds; gene; alpha 7 subunit; human;		
KW	KW	neuronal nicotinic acetylcholine receptor; ligand-gated ion channel;		
KW	KW	synaptic transmission; gene therapy; transgenic.		
XX	XX	Homo sapiens.		
OS	OS	Location/Qualifiers		
XX	XX	1. .72		
FH	FT	/*tag= a		
FT	FT	73.1584		
FT	FT	/*tag= b		
FT	FT	/product= "nAChR alpha 7 subunit"		
FT	FT	1585.1876		
FT	FT	/*tag= c		
XX	XX	US6524789-B1.		
PN	PN	25-FEB-2003.		
PD	PD	07-JUN-1996; 96US-00660451.		
XX	XX	07-JUN-1995; 95US-00484722.		
XX	XX	(MERI) MERCK & CO INC.		

QY	515	GC	GAGATGAAGTTTGGCAGCTTGGACTTATGATGGTTATACAGTTGGATCTCTAACTACAGG	574
Db	491	GC	AACTGAAGTTTGGTCTCGTCTTACGGAGGCTGGTCTTGGATCTGCAGATGCAGG	550
QY	575	AT	GAGGGGCGGAGATATAAGCAGTTTCTCAGAAATGCGAATGGAGTCTTAATAGGAG	634
Db	551	A	-----GGCAGATATCAGTGGCTATATCCCAATGGAGAAATGGGACTTGTGGGAA	601
QY	635	TC	CCCGCAAGCGAACAGAGATCTACTACAACTGTGTGTCGGAGCCATACATCAGACATCA	694
Db	602	TC	CCCGCAAGAGAGTGAAGGTTCTATCAGTGTCTCAAAAGAGCCCTACCCGATGTCA	661
QY	695	CG	TTTCGGGTGTATCCCGGAGGAACGCTCTACTACTTCTTCAATCTCATCGTGCCT	754
Db	662	CT	TTCACAGTGACCATCGCCCGAGGACGCTCTACTATGGCCTCAACCTGCTCATCCCT	721
QY	755	GC	GTGTCATCTCGCTCCATGGCTCTATTTGGGGTTCACTTTGCCTCCAGACTCCGGAGAAA	814
Db	722	GT	GTGTCATCTCCGCCCTCGCTCGTGTCTCTCTCTCTCTGTCAGATTCCTGGGAGA	781
QY	815	AG	TGTCTTTAGGTGACGATATTACTGTCTGTGACGGTGTCTCTCAATGGTGGCGG	874
Db	782	AG	ATTTCCCTGGGGATAACAGTCTTACTCTCTTACCGTCTTCATGCTCTCTGGCTG	841
QY	875	AG	ACGATGCCAGGACGCTCCGAGCCCGTGCCTTGTCTCGCACCTACTCTCAACTGATCA	934
Db	842	AG	ATCATGTCGCGCAACATCCGATTCGGTACCAATGATAGCCAGTACTTCGCCAGACCA	901
QY	935	TG	TTCATGTGTGGCTTCTCGTGTCTCCACCATATCTGATCTCTCAACTACACACCGCGC	994
Db	902	TG	ATCATCTGTGGGCTCTCGTGTGTGTGACGGTGATCTGTCTGTGACGACCAACCCACG	961
QY	995	ACC	CAGACACTCAGAAATCAGTATTGGATTTGTTGCGGTGTTCTTTATTGGTGCCTG	1054
Db	962	ACC	CCGACGGGGCAGATGCCAATGTGACACGAGTCTCTTCTGAACTGTGTGCGCGT	1021
QY	1055	GG	TGTGTGCATATGTCAACGCCCGGCTCGCGCAGCAGCCGCGCGCGCGGTACTCTC	1114
Db	1022	GG	TCTCTGCGAATGAAGAGGCCCGGGAGACAAAGTGTGCGCCGCGCTGCACACAAGC	1081
QY	1115	CG	CGC-----CGGACCTGGAGTGTGCGAGCGCTCCTCCAAGTGCCTCTAGCGA	1165
Db	1082	AG	CGCGCTCGACCTGTGGCCAGTGTGAGATGAGCGCGTGTGGCGCGCCGCCCGCAGCA	1141
QY	1166	AC	GTGTCGACATCGATGACACTTCGCGCACCCGCAAGCGCAGCAGCCCAATGTGTCC	1225
Db	1142	AC	GGAACTCTGTATCATCGGCTTCGCGGCTTCGACGCGTGGACGCGTGTGCCGACCC	1201
QY	1226	GAT	ACTACAGGGGGGTGAGAGNATGGCGGGGTGGCGGCGCAGATT-----	1276
Db	1202	CC	GACTCTGGGTGTGTGTGGCCGATGCGCTCTCCCAACAGCATGAGCACTCC	1261
QY	1277	-----	GCTTCGGTGTCACTACGAGCTCTCCCTCATTTCTTCAAGGAGATTA	1321
Db	1262	TG	CAGGTGGGCAACCCCGAGGGGACCGGACTTGGCCAGATCTCTGGAGAGGTCC	1321
QY	1322	GAG	TATCAGATCTAGATGCGCAGAGAGATGCGGACATTTTCGGCGGACTGGA	1381
Db	1322	GCT	ATATTGCAACCGCTTCGCTGCCAGACCAAGCAGGGGGTCTGCAGCGAGTGGGA	1381
QY	1382	AG	TTCGCGGCATGTGTGTGACAGACTGTGCTTATTATCTTTACCTCTTCCAAATCA	1441
Db	1382	AG	TTCGCGCTGTGTGTGACCGCTGTGCTCTCATGGCTCTCTCGTCTTACCAATCA	1441
QY	1442	TC	GCACGCTAGCCGTGCTGTCTCGCGCAACACATCATGTGGTGTCT	1487
Db	1442	TC	TGCACCATGGCATCTCATGTCTGGCTCCCACTTCGAGAGGC	1487

XX	AAC90385;
XX	
XX	14-MAR-2001 (first entry)
XX	
XX	Mutant human alpha7 ligand gated ion channel coding sequence #1.
XX	
XX	Human; alpha7 nicotinic acetylcholine gated ion channel; mutant;
KW	5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200073431-A2.
XX	
PD	07-DEC-2000.
XX	
XX	25-MAY-2000; 2000WO-US011862.
XX	
PR	27-MAY-1999; 99US-0136174P.
XX	
PA	(PHAA) PHARMACIA & UPJOHN CO.
XX	
PI	Groppi VE, Wolfe ML, Berkenpas MB;
XX	
DR	WPI; 2001-061524/07.
DR	P-PSDB; AAB50015.
XX	
PT	Special cell culture medium for treating cells and for inducing mammalian
PT	cell lines to conduct calcium ions, comprising specified concentrations
PT	of ions of sodium, calcium and potassium at specified pH.
XX	
PS	Claim 50; Page 69; 77pp; English.

		251	ATCAGTATTACAGTGGAAATGTGTCAGAAATATCCAGGGTGGAAGACTGTTCTGTTTCCAG	310
Db				
		335	CCCACAGCATATGAAAACCAGACGTCCTTATGTACAACACGGCGACCAAGGGTTCGACA	394
Qy				
		311	ATGCCAGATTGGAAACCAGACATCTTCTCTATACAGTCTGTATGATGGCGCTTTGACG	370
Db				
		395	GACGTAATCCAACGAACGTGGTGGTCGGGAACAAACGGCTCGTGTCTTACGTGCACGGCCG	454
Qy				
		371	CCAATTTCCACACTAAGCGTGTGTGTAATCTTCTGGGCATTCGCAGTACTGCTCTCCAG	430
Db				
		455	GCATCTTCAAAGACACCTGCAAGATGCACATCACCTCGTTCCCTTCGACGACCAACGAT	514
Qy				
		431	GCATATTCAGAGTTCCTGCTACATCGATGTACGCTGTTTTCCCTTGATGTGCAGCACT	490
Db				
		515	GCGAGATGAAGTTTGGCAGCTGGACTTATGATGGTTATCAGTTGGATCTCAACTACAGG	574
Qy				
		491	GCAACTGAAGTTTGGCTCGTGGTCTTACGGAGGCTGGTCTTGGATCTGCAGATGCAGG	550
Db				
		575	ATGAAGGGGGCGGAGATAAAGCAGTTTGTGACGAATGGCAATGGAGTTAATAGGAG	634
Qy				
		551	A-----GGCAGATATCAGTGGCTATATCCCAATGGAGAATGGGACCTAGTGGGAA	601
Db				
		635	TCCCGGCAAGCGCAACGAGATCTACTACAACTTTGTTCGGAGCCATACATCGACATCA	694
Qy				
		602	TCCCGGCAAGAGGAGTGAAGGTTCTATAGTGTCTCAAAGAGCCCTACCCCGATGTCA	661
Db				
		695	CGTTTGGGTGGTGATCCGGAGGAAAGCTCTACTACTTCTTCAATCTCATCGTGCCT	754
Qy				
		662	CCTTTCAGGTGACCATCGCGCGAGGACGTCTACTATGGCTCAACCTCTGATCCCCA	721
Db				
		755	GCGTGCTCATCGCCTCCATGGCTCTATTTGGGGTTCACTTTCCTCCAGACTCCGGAGAAA	814
Qy				
		722	GTGTGCTACTCGCCCTCGCCCTGTGGTGTCTTCTTCCTGCAGATTCCGGGAGA	781
Db				
		815	AGTTGCTTTTAGTGTGCAGANATTAATGTGCTTGAACGGTTCCTCAACATGSGTGGCG	874
Qy				
		782	AGATTTCCCTGGGGATAACAGTCTTACTCTCTTACCGTCTTCATGCTGCTGTGGCTG	841
Db				
		875	AGACGATGCCAGCGACGTCGGAGCGCTGCCCTTGTCTGGCACCTACTTCAACTGCATCA	934
Qy				
		842	AGATCATGCCCGAAACATCGATTTCGTAACATTGATAGCCAGTACTTTCGCAGACCA	901
Db				
		935	TGTTCAATGTGGCTTCTCCGTGTCTCCACCATACTGATCTCTCAACTCACCAACCGGC	994
Qy				
		902	TGATCATGTGGGCTCTCGGTGGTGTGACGGTGATCGTGTCTTACCGTCTTCATGCTGTG	961
Db				
		995	ACGCAGACACTCAGAAATGAGTGAATGGATTCGTTGGTGTTCCTTTATTTGCTGCGCT	1054
Qy				
		962	ACCCGACGCGGGGCAAGATGCCAAGTGGACCAAGTGCGCCGCTGCCAGCACAAAGC	1021
Db				
		1055	GGGTGCTTCGCATGTACGGCCCCGGCTCGGCGACGACCGCCGCCGCGCGGTACTCTC	1114
Qy				
		1022	GGTTCTTGCGAATGAAGAGGCCCGGGGAGACAAGGTGCGCCGCTGCCAGCACAAAGC	1081
Db				
		1115	CGCG-----CCGGA CCTGSAGCTGGCGAGCGCTCCTCAAAGTGGCTCTTAGCGA	1165
Qy				
		1082	AGCGGCGCTGCAGCTTGCGCAGTGTGAGATGAGCGCGGTGGCGCCGCCGCCCGCAGCA	1141
Db				
		1166	ACGTGCTCGACATCGATGACGACTTCGCGCACCCCGCAAGCGCAGCGCGCAATGCTGCC	1225
Qy				
		1142	ACGGGAACCTGCTGTACATCGGCTTCGCGGCTTGACGGCGTGCACTGTCTCCGACCC	1201
Db				
		1226	GATTA CTACAGGGGGGTGAGAGAAATGGCGGGGGTTGGCGGCGCA CAGTTGCTTGGTG	1285
Qy				
		1202	CCGACTCTGGGGTAGTGTGTGGCGCGATGGCTCTCTCCCA CGACGATGAGCACCTCC	1261
Db				
		1286	TCGACTACGAGCTCTCCCTC-----ATTCTGAAGGAGATTA	1321
Qy				
		1262	TGCACGGCGGCAACCCCGAGGGGACCCGAGCTTTGGCCAAGATCTTGAGGAGGTCC	1321
Db				
		1322	GAGTCATCACAGATCAGATCGCAAGGACGACGAAGATGCGGACATTTCCGCGCACTGGA	1381
Qy				

[illegible]

Db 131 TGCCCAATGACTCGCAACCACTACCGCTCTACTTCTCCCTGAGCCTCCTGAGATATGG 190
QY 215 ACGTGGACGAGAAACCAAGCTTTTAAATAACAAACATCTGGCTAAACTAGAGTGGAAATG 274
Db 191 ACGTGGATGAGAAACCAAGCTTTTAAACCAACCAACATTTGGCTGCAAAATGCTTTGGACAG 250
QY 275 ATATGACTTGGGTGGAACACTTCAGATTTTCGGCGGGGTCAAGATTTAAGATGGCCAC 334
Db 251 ATCACTATTTACAGTGAATGTGTCAGAAATATCAGGGGTGAAGACTGTTCGTTTCCAG 310
QY 335 CCACAGACTATGGAACACCAAGACGCTCTCTATGTACAACAGCGGACGAAAGGTTTCGACA 394
Db 311 ATGSCCAGATTTGGAACACCAAGACATTTCTCTATACAGTGTGATGAGCGCTTTGACG 370
QY 395 GCACGTATCCAAAGACGCTGGTGTGCGGAACAAACCGCTCGTGTCTGTACGTGCGGCCCG 454
Db 371 CCACATCCACACTAAACGCTGTGGTGAATTTCTCTGGGCATTTGCCAGTACCTGCTCCAG 430
QY 455 GCATCTTCAAGACACCTCGAAGATCGACATCACCTGGTTCCCTTCGACGACCAAGAT 514
Db 431 GCATATTCAGAGTTCCTGCTCATCGATGTACGCTGGTTTCCCTTTGATGTGAGCACT 490
QY 515 GCAGATGAATGTTGGCAGCTGGACTTATGATGGTTATCAGTTGGATCTACAATACAGG 574
Db 491 GCAACTGTAAGTTTGGGTCTGGTCTTACGAGGCTGGTCTTGGATCTCGAGATGCAGG 550
QY 575 ATGAAGGGGGGAGATATGAAGAGATTTTCTACGAATGCGAATGGAGTTAATAGGAG 634
Db 551 A-----GGCAGATATCAGTGGCTATATCCCAATGAGAATGGACCTAGTGGGAA 601
QY 635 TCCCGGCGAGCGCAAGAGATCTACTACACTGTTGTCCGGAGCCATACATCGACATCA 694
Db 602 TCCCGGCGAAGAGTGAAGGTTCTATGAGTGTCTGCAAGAGCCCTACCCCGATGTCA 661
QY 695 CGTTTGGGTGGTGATCCGGAGGAAACCGCTCTACTACTTCTTCAATCTGATCGTGCCCT 754
Db 662 CTTTCAGATGACATGCGCGCGAGCGCTCTACTATGGCTCAACTGCTGATCCCCA 721
QY 755 GCGTGCTCATCGCTCCAGGCTCTATTGGGGTTACCTTGGCTCCAGACTCCGGAGAAA 814
Db 722 GTGTGCTCATCTCGGCTCGGCTGCTGGTGTCTCTGCTTCTGTCAGAGATTCGGGGAGA 781
QY 815 AGTTGCTTTAGGTGTGACATATTACTGTGTCAGGTGTTCTCAACATGGTGGCGG 874
Db 782 AGATTTCCCTGGGATTAACAGTCTTACTCTCTTACCGTCTTCAATGCTGCTGCTG 841
QY 875 AGACGATGCCAGGACGCTCGGAGCGCGTGGCTTGTCTCGGCACCTACTTCAACTGCATCA 934
Db 842 AGATCATGCCCGCAACATCCGATTCGGTACCATTTGATAGCCAGTACTTCGCCAGACCA 901
QY 935 TGTTCATGGTGGCTTCTCGTGTCTCCACCATATGATTCCTCAACTACCAACACGGG 994
Db 902 TGATCATCGTGGGCTCTCGGTGGTGGTGAAGCGGTGATCGTGTGTCAGTACCAACACG 961
QY 995 ACGCAGACACTACGAAATGAGTGGATTGTTGCGGTGTTCTTTATTGGCTGCCGT 1054
Db 962 ACCCGACGGGGCAAGATGCCAAGTGGACCAAGTCACTCTCTGAACTGGTGGCGGT 1021
QY 1055 GGGTGCTGCGCATGTCAACGCCCGGCTCGGCGACGACGCGCGCGCGCGCGGTACCTC 1114
Db 1022 GGTTCCTGCGAATGAAGAGCCCGGGAGGACCAAGTGGCGCCCGGCTGCGCAGCAAGC 1081
QY 1115 CGCG-----CGGACCTGGAGCTGCGAGCGCTCTCTCAAGTGGCTCTTAGCGA 1165
Db 1082 AGCGGCGCTGACGCTGGCAGTGTGAGATGAGCGCGGTGGCGCGCGCGCGCGAGCA 1141
QY 1166 ACGTGTTCGACATCGATGACGATTCGCGCCACCGCAAGCGCAGCAGCGCAATGTGCC 1225
Db 1142 ACGGGAACCTGTGTATCGGCTTCGCGGCTTGACGCGCTGCACTGTGTCCGACCC 1201
QY 1226 GATACTACAGGGGGGTGAGGAGAAATGGCGCGGGTTGGCGGCGCACAGTTGCTTGGTG 1285

Db 1202 CCGACTCTGGGGTAGTGTGTGGCGCGCATGCGCTGCTCCCCACGCCACGATGAGCACCTCC 1261
QY 1286 TCGACTACGAGCTCTCCCTC-----ATTCTGAAGGAGATTA 1321
Db 1262 TGCACGCGCGGCAACCCCGAGGGGACCCCGGACTTGGCCAAGATCTTGGAGGAGTCC 1321
QY 1322 GAGTCATCAGATCAGATCGGCAAGGACGACGAAGATGCGGACATTTCCGGGAGCTGGA 1381
Db 1322 GCTACATTTGCCAATCGCTTCCGCTGCCAGGACGAAAGCGAGGCGGTCTGCAGCGAGTGA 1381
QY 1382 AGTTGCGCGCATGCTGTCGACAGACTGTGCTTATTATCTTTACCTCTTCAACAATCA 1441
Db 1382 AGTTGCGCGCTGTGTGTGGACGCGCTGTGCTCTATGGCTTCTCGGTCTTCAACATCA 1441
QY 1442 TCGCCACGCTAGCGGTGCTGCTGTCCGCGCCACACATCATGCTGTC 1487
Db 1442 TCTGCACCATCGGCATCTGATGTGCGCTCCCAACTTCGTGGAGC 1487

Search completed: May 7, 2004, 15:01:41
Job time : 609.456 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:08:32 ; Search time 3998.88 Seconds

(without alignments)
11111.850 Million cell updates/sec

Title: US-09-303-232-3_COPY_335_1822

Perfect score: 1488

Sequence: 1 atggcgggcgggcgcgcg.....cgccacacatcatgtgtcg 1488

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_esti:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estcom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	476.2	32.0	885	12	BG632919
2	320.6	21.5	1201	9	AL530299
3	309.2	20.8	2296	14	CD013901
4	281.6	18.9	1864	11	AK053497

5	281.6	18.9	2916	11	AK051730	Mus muscu
6	281.6	18.9	3126	11	AK080415	Mus muscu
7	280.4	18.8	4290	11	AK029177	Mus muscu
8	278.4	18.7	1374	29	AY406232	Mus muscu
9	275.2	18.5	1374	29	AY406230	Mus muscu
10	274.2	18.4	1466	14	CD013891	Homo sapi
11	271.4	18.2	1436	29	AY402873	Homo sapi
12	269.6	18.1	1436	29	AY402875	Mus muscu
13	269.4	18.1	607	9	AI292581	GH1518.5
14	263.4	17.7	1436	29	AY402874	Pan trogl
15	263.4	17.7	2010	11	AK080475	Mus muscu
16	261.8	17.6	1374	29	AY406231	Pan trogl
17	258	17.3	4037	11	AK049722	Mus muscu
18	258	17.3	4046	11	AK051742	Mus muscu
19	253.8	17.1	2940	11	AK034228	Mus muscu
20	253.8	17.1	3230	11	AK083157	Mus muscu
21	245.2	16.5	908	13	BUI49265	AGENCOURT
22	244.4	16.4	3483	11	AK081254	Mus muscu
23	241.4	16.2	615	14	CB149460	K-EST0205
24	239.8	16.1	1442	29	AY402876	Homo sapi
25	236.8	15.9	1034	13	BX403124	EX403124
26	232.8	15.6	922	13	BUI15857	AGENCOURT
27	226.2	15.2	411	12	B1516733	BB160023A
28	226	15.2	1442	29	AY402877	Pan trogl
29	225.2	15.1	410	12	B1516843	BB160023B
30	222.6	15.0	1454	29	AY402878	Mus muscu
31	216	14.5	902	29	AY407186	Mus muscu
32	214.8	14.4	1603	29	AY411327	Homo sapi
33	214.6	14.4	833	14	CB245337	UI-M-FY0-
34	213.8	14.4	640	14	CB244439	UI-M-FY0-
35	212.8	14.4	1101	29	CNS0001F	Drosophil
36	212.8	14.3	1007	29	CNS000HUU	Drosophil
37	208.6	14.0	1781	14	CD013889	90138814
38	208	14.0	1036	13	BX437801	BX437801
39	203.8	13.7	978	29	CNS006F9	Drosophil
40	202	13.6	720	14	CA373069	647053 NC
41	199.6	13.4	2513	11	AK033068	Mus muscu
42	198	13.3	755	10	AW914206	EST345510
43	197.2	13.3	791	14	CF739825	UI-M-HD0-
44	191.6	12.9	864	9	AUI20692	AUI20692
45	189.6	12.7	658	12	BM711715	UI-E-CL1-

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS
DEFINITION
BG632919 885 bp mRNA linear EST 23-APR-2001
GH16126.3prime GH Drosophila melanogaster head pOT2 Drosophila
melanogaster cDNA clone GH16126 3 similar to CG4128; Fban0004128
'ion channel', located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BG632919
BG632919.1 GI:13758409
EST
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidae; Drosophilidae; Drosophila.
1 (bases 1 to 885)
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G.M.
BDGP/HMMI Drosophila EST Project
Unpublished (2001)
Other ESTs: GH16126.5prime
Contact: Stapleton, N.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or

more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003511: arm.X [18792641..19136447] estimated-cyto:18A3-18C6: 04/10/2001 Plate: GH.161 row: C column: 2 High quality sequence stop: 784.

FEATURES

Location/Qualifiers
1..885
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH16126"
/sex="male and female"
/dev_stage="adult"
/lab_hosts="DHS - alpha"
/clone_lib="GH Drosophila melanogaster head pot2"
/note="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 32.0%; Score 476.2; DB 12; Length 885;
Best Local Similarity 73.2%; Pred. No. 8.9e-82;
Matches 610; Conservative 0; Mismatches 223; Indels 0; Gaps 0;
Qy 193 CTCACGCTCATGCGAGATCATCGACGTGGACGAGAGAACACGCTTTAAATAACAAACATC 252
Db 885 CTACACTCATGAGATATTCGATGTGACGAGAGATCACTGCTTTATACGAATATT 826
Qy 253 TGGCTAAACTAGAGTGGATGATPATGAACCTTGAGTGGACACTTCAGATTTCCGGCGG 312
Db 825 TGGCTCAATTGGAATGAACGATATGAATCTTCGATGGAATTCGAGTGGATTCGGTGGT 766
Qy 313 GTCAAGATTTAAGAGTGCACCCACAGACTATGGAACAGACAGCGTCTTATGTACAAC 372
Db 765 GTGCGGATCTGCGAATTCGCGCCACATCGCCCTATGGAACCGGATGACTGATGTACAAC 706
Qy 373 AGCGCGACGAAGGTTTCGACGACACGCTATCCAAACGAGCTGGTGGCGGAACAACGGC 432
Db 705 AGTCCGACGAGGCTTCGATGGACGATACGCCACAATGTGGTTCGCAATATGGG 646
Qy 433 TCGGTGTGATGTCGCGCCCGCGGATCTTCAAGAGCACCTGCAAGATCGACATCACTGG 492
Db 645 AGCTGTCTGACGTACGCGCCAGGTATATTAAAGTCAACGTGAAGATCGACATTCAGTGG 586
Qy 493 TTCCCTTCGACGACCAACGATCGAGATGAAGTTTGGCAGCTGGACTTATGATGGTAT 552
Db 585 TTTCATTCGACGATCAGATGTGAATGAAATTTGGTTCGTGGACCTAGATGGGTTT 526
Qy 553 CAGTTGGATCTACAACCTACAGGATGAAGGGGGCGGAGATATAAGCAGTTTTGTCAACGAAT 612
Db 525 CAGTTGACCTGCAGTTTCGAGGACGAAGCTGGTGGCGACATTTCTAGCTTTATACCAAT 466
Qy 613 GCGAATGGAGATTAAAGAGTCCCGGACGAGCGCAACGAGATCTACTACAATGTTGT 672
Db 465 GCGAATGGAGCTTGTAGGTGTGCGCGGTAAACGAAATCTACTATAATTTGCTGC 406
Qy 673 CCGGAGCATACATCGACATCACGTTTGGGTGTGATCCGGAGGAAACGCTACTATAC 732
Db 405 CCAGAACCTTATATGACATAACATTCGCCATTTTGAATGAGCGCAAAACGTTTACTAT 346
Qy 733 TTCTCAATCTGATCGTCCCTGCGTCTCATCGCTCCATCGGCTCTATTGGGGTTCAAC 792
Db 345 TTTTCAATCTGATTTGCGCGTGGTACTGATCGCTCCATCGGCACTGCTAGGGTTTACA 286
Qy 793 TTGCTCCAGACTCGGAGAAAAGTTGTTTAGGTGTGAGATTAATTAATGCTGTTGACG 852
Db 285 CTGCCACAGATTTCTGGTGAAGACTTTTCGTTGGAGTTTCAATTTCTATTATCGCTTACA 226
Qy 853 GTGTTCTCAACATGGTGGCGAGACGATGCCAGCGAGCTGGACGCGCTGCCCTTGC 912
Db 225 GTCTTCTCAACATGGTGGCGAACAATGCCGCGACCTCCGATGGGTACCGCTGCTC 166

Qy 913 GGACCTACTTCAACTGCATCATGTTGCTGCTCTCCGCTGCTCCACATACCTG 972
Db 165 GGAATTTATTTCAATTCATATGTTTATGGTGGCTCATCAGTTGTGCAACATATT 106
Qy 973 ATCTCAACTACCAACCGCGACGAGACACTCACGAAATGAGTATTGGAT 1025
Db 105 GTCCTCAATTATCATCATAGAAATCCAGATACGATGAATGAGTGAATGGT 53
RESULT 2
AL530299 1201 bp mRNA linear EST 23-MAY-2003
LOCUS AL530299 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DD007YP05 5-PRIME, mRNA sequence.
ACCESSION AL530299
VERSION AL530299.2 GI:31068132
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:12793792.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7646.r For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD007CH03QP1&cluster=7646.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DD007CH03QP1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD007YP05"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 21.5%; Score 320.6; DB 9; Length 1201;
Best Local Similarity 60.0%; Pred. No. 1.4e-51;
Matches 564; Conservative 5; Mismatches 360; Indels 11; Gaps 2;
Qy 9 GCGGGCGCGCGCTCGCACTTGGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCT 68
Db 108 GCAGTCCGGAGTCAACATCGCTGCTGCGGGAGGTGCTGCTGCTGCTGCTGCTGCT 167
Qy 69 CTGCGCGAGGGGGGACGCTGCGGGTACACGAGAGCGGCTACTGCACCACTTATGGA 128
Db 168 CGTCTCTGACGTGCTCCCTGCAACGAGTWCAGAGGAAGCTTTACAAGAGCTGGTCAA 227
Qy 129 CCAGTCAACGCTACTGGAGAGGCGCGTCTGCTCAAGAGAGCGACCGCTGAGCTCTCTT 188
Db 228 GAATCAATCCCTTGGAGAGCGCGCGTGGCCAAATGACTCGCAACCACTACCGTCTACTT 287
Qy 189 CGGCTCACGCTCATGCGAGATCATCGACGTGGAGAGAGAACAGCTTTTAAATAACAAA 248
Db 288 CTCCTGAGCTCTCTGAGATCATGACGTGGATGAGAGAACCAAGTTTAAACACCAA 347
Qy 249 CATCTGCTCAAACTAGAGTGAATGATATGAATCTGAGGTGGAACACTTCAGATTTGCG 308


```
Db 348 CATTGGCTCAAAATGCTTTGGACAGATCACTATTTACAGTGGATGTGTCAGAAATACC 407
QY 309 CGGGGTCAAAGATTTAAGAGTGCACCCACAGACTATGGAACACAGACGTCTCTATGTA 368
Db 408 AGGGGTGAAGACTGTGTTCCAGATGCCAGATTTTGAACACAGACATCTCTCTCTA 467
QY 369 CAACAGCGCGGAGGAAGGTTGCAGACAGCTATCCAAACAGTGGTGGTGGGAACAA 428
Db 468 TAAACAGTGTGATGAGCGCTTTGACGCCACATTCACACATTAACGCTGTGTGTAATCTTC 527
QY 429 CGGCTCGTGTCTGTACGTGCGCGCGCATCTTCAAGAGCACCTGCAAGATCGACATCAC 488
Db 528 TGGGCATTGCGAGTCTGCTCCAGGCATATTCAGAGTCTCTGCTACATCATGATGAC 587
QY 489 CTGGTTCCCTTCGAGACCAACGATCGAGATGAAGTTTGGCAGCTGGACTTATGATGG 548
Db 588 CTGGTTCCCTTTGATGTGCAGCACTGCAAACTGAACTTTGGTCTCTGGTCTTACGGAGG 647
QY 549 TTATCAGTTGGATCTACACTACAGATGAAGGGGGGGAGATATAGCAGTGTTCAC 608
Db 648 CTGGTCTCTGGATCTGCAGATCGAGGAGGC-----AGATATCAGTGGCTATATCCC 698
QY 609 GAATGGCGAATGGAGTTAATAGAGTCCCGGCAAGCGCAACGAGATCTACTACAACCTG 668
Db 699 CAATGGAGATGGACCTAGTGGATCCCGCAAGAGAGTGAAGGTTCTATGAGTK 758
QY 669 TTGTCGGAGCCATACATCGACATCACGTTTGGCGTGGTGTATCCGGAGAAACGCTCTA 728
Db 759 CTGCAAGAGCCCTACCCCGATGTCACCTTCAAGTGAACCATGCGCGCAGGACGCTCTA 818
QY 729 CTACTTCTCAATCTGATCGCTCGCTGCTCATCGCTCCATGCTCATGCTCTATGGGTT 788
Db 819 CTATGGCTCAACCTGCTGATCCCTGTGCTCATCTCCGCCCTCGCCCTGCTGGTGT 878
QY 789 CACCTTCCCTCCAGACTCCCGAG--AAAAGTTGCTTTAGTGTGACGATATTACTCTCG 846
Db 879 CTTGCTCTCTCGAGATTCGGGGGAGAGATTTCCCTKGGGATAACAGTCTTACTCTCT 938
QY 847 TTGACGGTGTCTCTCAACATGGTGGCGGAGACGATGCCAGCGTCGCGCGCGTGC 906
Db 939 CTTACCGTCTCTATGCTGCTGCTGCTGATCATGATGCGCGCAACATCCGATTCCGTACCA 998
QY 907 TTGCTCGGCACTTCTCACTGCATCATGCTCATGCTGG 946
Db 999 TTGAAGCCCGAGTCTTTCGGCAGCACCATGATCATCGTG 1038

RESULT 3
LOCUS CD013901
DEFINITION 90134548 single gene library Homo sapiens cdNA, mRNA sequence.
ACCESSION CD013901
VERSION CD013901.1 GI:37777431
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2296)
AUTHORS Jin,P., Fu,G.K., Willson,A.D., Yang,J., Chien,D., Hawkins,P.R.,
Au-Young,J. and Stuve,J.L.
TITLE PCR isolation and cloning of novel splice variant mRNAs from known
drug target genes
JOURNAL Unpublished (2003)
COMMENT Contact: Jin, P.
Incyte Corporation
3160 Porter Drive, Palo Alto, CA 94304, USA
Tel: 650 621 8639
Fax: 650 621 8965
Email: pj@incyte.com.
FEATURES
source Location/Qualifiers
1..2296
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/notes="Vector: pDrive Cloning Vector; RT-PCR was performed using gene-specific primers flanking the open-reading frame. PCR products were subcloned into pDrive Cloning Vector and sequenced completely using M13 forward and reverse primers. Sequencing gaps were closed by re-sequencing using primers flanking the gapped areas."

ORIGIN

Query Match 20.8%; Score 309.2; DB 14; Length 2296;
Best Local Similarity 58.0%; Pred. No. 2.8e-49;
Matches 567; Conservative 0; Mismatches 408; Indels 3; Gaps 1;

QY 100 GAGAAGCGGTACTGCACCACCTATTGGACCACTACAAACGTACTGGAGAGCGCCGTCGTC 159
Db 350 GAGGACGGGTCTTCAAACACCTCTCCGGGGCTACAAACGCTGGGGCGCCCGGTGCC 409
QY 160 AACGAGAGCGACCGCTGCGAGCTCTCTTCGGGCTCACGCTCATGAGATCATCGAGTG 219
Db 410 AACATTCAGACGCTGTGATGTGGCTGTGGACTGTCCATGCTCAGCTCATCATGATG 469
QY 220 GACGAGAGAACGAGCTTTTAATAACAACATCTGGCTAAACTAGAGTGAATGATG 279
Db 470 GATGAGAGAACCAAAATGATGACCAACGCTGTGGCTAAACAGGAGTGGAGCATAC 529
QY 280 AACTTCAGGTGGAAACATTCAGATTTTCGGGGGTCAAAAGATTTAAGAGTGCACCCAC 339
Db 530 AACTTCGGTGGAAACCCGCTGATTTGGCAACATCACATCTCTCAGGTCCTCTCTGAG 589
QY 340 AGACTATGGAACACAGAGCTCTTATGTACAACAGCGCGGACGAAGGTTGACAGACG 399
Db 590 ATGATCTGATCCCGACATTTGTTCTTACAACATGAGTGGGAGTTTGCAGTGACC 649
QY 400 TATCCAGCAAGCTGGTGGTGGGACCAACGCTCGTGTCTGTACGTCGCGCGGATC 459
Db 650 CACATGACCAAGGCCACCTCTTCTCCACGGGCACTGTGCACCTGGGTGCCCCGGCCATC 709
QY 460 TTCAAGAGCACTTCAAGATCGACATCACCTGGTTCCCTTTCGACGACCAACAGATCGAG 519
Db 710 TACAAGAGCTCTTCGACGATCGAGCTGCTGCGCGGAGATCTACCCGACGTCACCTAC 946
QY 520 ATGAAGTTTGGCAGCTGGAATATGATGATGATGATGATGATGATGATGATGATGATG 579
Db 770 ATGAAGTTTGGCTCTCGACTTATGACAAAGGCCAAGATCGACCTGGA--GCAGATGGAG 826
QY 580 GGGGGGGAGATATAAGCAGTCTTGTCAAGAAATGGGGAATGGGAGTTAATAGGATCCCC 639
Db 827 CAGACTGTGGACCTGAGGACTACTGGGAGAGCGGGAGTGGGCCATCTGTCATATGCCACG 886
QY 640 GGCAAGCGCAACAGAGATCTACTACACTGTGTTCGGAGCCATACATTCGACATCACCTTT 699
Db 887 GGCACCTTACAACAGCAAGAGTACGACTGCTGCGCGGAGATCTACCCGACGTCACCTAC 946
QY 700 GCGGTGGTATTCGGAGGAAAACGCTCTACTACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 759
Db 947 GCCTTGTCTATCGGGGGCTGCGGCTCTTCTACACCATCAACCTCATCATCCCTGCTG 1006
QY 760 CTGATCGCTCCATGCTCTATTTGGGGTTCACCTTGCCTCCAGACTCCGAGAAAGTTG 819
Db 1007 CTCATCTCCGCTCTCATGTGTGGTCTTCTACCTGCCCTCCGAGTGGCGGAGAGATC 1066
QY 820 TCTTTAGGTGTGACGATATTACTGTGTTGACGCTGTCTCTCAACATGGTGGCGGAGACG 879
Db 1067 ACGTGTGCATTTGGGTGTGCTGTCACTACCGCTCTTCTGCTGTCTCATCTGAGATC 1126
QY 880 ATGCCAGCGAGCTCGGACGCCGCTGCTCGGACCTTACTTCAACATCGATCATGTTTC 939
Db 1127 ATCCCGTCCACCTTCGCTGGTCTATCCCGCTCATCGGCGAGTACCTGCTGTTCACCATGATC 1186
QY 940 ATGGTGGCTTCTCCGCTCGTCTCCACCATACTCTCACTACACACCGGACGCA 999
```

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resesgc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

URL: <http://rancom.gsc.rken.go.jp/> .

Location/Qualifiers
1. .1864

```

strain="C57BL/6J"
db_xref="PANTOM.DB.E130103E14"
db_xref="MGI:243283"
db_xref="taxon:10090"
clone="E130103E14"
tissue_type="eyeball"
clone_lib="RIKEN full-length enriched mouse cDNA library"
dev_stage="0 day neonate"

```

```

94. .1593
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_size="0 day neonate"
/note="unnamed protein product; NEURONAL NICOTINIC
ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT (SPTR)AAL58471,
evidence: FASTV, 99.8%ID, 100%length, match=1497)
outative"

```

/protein_id="BAC35404.1"
/db_xref="GI:26343495"
translation="MGVVLPPPLMLMLVLPVASAEAEHLFOYLPEDYNELI
RPNVASHPVLIQEVSMGSLVKDEUNMETNLMLKQNDYLLKKWSPDQGVGF
MRVPAEKIKWEDIYLNNAADQFQDDTKALKLTKIGVTWIPALFASCKUDIVF
PFIQYCNCTMFGSGWDRADKDLVLIGSSMMLQDWESGELLAKGPKYKEIKVNC
EEIYQDITDITPSTLPIVLEGYLLFTMTLFSIVTVFVNLHVHRTTTHMPT
LVFLVLTITPSTLPIVLEGYLLFTMTLFSIVTVFVNLHVHRTTTHMPT
WKVAFVNLVPMVMTSTEDAKTRFNFGAELSNLCSFKRADSKSGVPCQ
QDQYCHHRRVTKSNFSANLSSRSESVDAVLSALSPEIKAEIOSVKYIAENMK
AGTVAKIEQDWKVAAMVDIRFLWVFLVLCVLTAGTFLQPLMARDOT"

```

1864      /note="putative"
1865      18.9%; Score 281.6; DB 11; Length 1864;
1866      Similarity 55.0%; Pred.No.5.8e-44;
1867      Conservative 0; Mismatches 463; Indels 3; Gaps 1;

```

TGGCGGCGCCGCGGGGCTGTCTGTCGTGTGCCCTCTGSCCGAGGGGGCACACT 88
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7 TGGCCCCGCGCCCGCTGTCCANGCTGATGCTGTGTCTGATGCTGTGCAGTGGCCACGG 166
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
19 GCGGGTACCAGGAAGCGGTACTTGCAACCACTATTGGACCATACAACTACTGGAGA 148

7 CCTCGAAGCTGAGCAGCGCTGTTTCCAGTACCTGTTTGAAGATTACAACGAGATCATCC 226
 9 GCGCGTGGTCAACGAGAGCGACCGGTGCAGCTCTCCTTCGGCCTCAGCTCATGCGAGA 208
 7 GCGCGTGGCTAAAGTGTGCCATCCTGTATCATCCAGTTTTCAGTGTCTATGTCTCAGC 286
 9 TCATCGAGCTGGACGAGAAGAACAGCTTTTAAATAACAACATCTGGCTAAAACCTAGAGT 268
 7 TGGTGAAGGTTGGATTGAAGTAAACAGAGATCATGGAACCAACCTGTGGCTGAAGCAAATCT 346

QY	269	GGATGATATGAACCTTGGAGTGGAACTTCAGATTTCGGCGGGGTCAAAGATTAAAG	328
Db	347	GGATGACTCAAGCTGAATGGAAACCTCTGACTACCAAGGGGTGGAGTTCTATGGCAG	406
QY	329	TGCCACCCACAGACTATGGAACCCAGACGTCCTTATGTACAAACAGCGCGGACGAAGGGT	388
Db	407	TCCCTGCAGAGAAGATCTGGAACCCAGACATCGTGCTTTACAAACAACGCCGATGGGATT	466
QY	389	TCGACAGCAGTATCCAAAGCAAGTGGTGGTCGGGAACAACGGCTGTGTCTGTACGTGC	448
Db	467	TCCAAGTGGATGACAAACCCAAAGCTCTACTCAAGTATCACAGGAGAAGTGACTTGGATCC	526
QY	449	CGCCCGGCATCTTCAAGAGACCTGCAAGATTCGACATCACCTGTGTCCCTTTCCAGCACC	508
Db	527	CTCGGCCAATCTTTAAGAGCTATGCAAAATCGATGTGACCTACTTCCCGTTTGACTTACC	586
QY	509	AACGATCGGATGAAGTTTGGCAGCTGGACTTATGATGGTATTACGTTGGATCTACAAC	568
Db	587	AAAACCTGCACCATGAAGTTGGCTCCTGGTCCTACGACAAGGCAAGATCGACCTGGTCC	646
QY	569	TACAGGATGAAGGGCGGAGATATAAGCAGTTTGTTCACGAATGGCGAATGGGAGTTAA	628
Db	647	TCA---TTGGCTCTCCATGAACTCAAGACTATTGGAAAGTGGCGAGTGGGCGCATCA	703
QY	629	TAGGAGTCCCGGCAAGCGCAACAGATCTACTACACTGTGTGTCGGGAGCATACATCG	688
Db	704	TTAAAGCCCGGGCTACAAACATGAATCAAGTACAACTGCTGTGAGGAGATCTACCAAG	763
QY	689	ACATCAGTTTGGCGTGGTGNATCGGAGGAAAACGCTCTACTACTTCTTCAATCTGATCG	748
Db	764	ACATCAGTACTCGCTATACATTCCGCGCCTCGCGCTGTTCTACCATCAACGCTCATCA	823
QY	749	TGCCCTCGCTGCTCATCGCTCCATGCTCTATTGGGTTTCACCTTGCTCCACGACTCCG	808
Db	824	TTCCGTGCTGCTCATCTCCTTCTCTACTGTGCTGTCTTCTACTGCCCCTCGACTGTG	883
QY	809	GAGAAAGTTGTCTTTTAGGTGTGACGATATPACTGTGTGTGACGGTGTTCCTCAACATGG	868
Db	884	GGGAGAAGGTGACGCTGTGCATCTCCGTGCTCCTCTCCCTGACCGTCTTTCTCCTCGTGA	943
QY	869	TGGCGGAGACGATGCCAGCAGCTCGGACGCGGTGCGCTTGTCTCGGACACTTCACTTCACT	928
Db	944	TCACCGAGACCATCCCTTCCACCTCACTGGTTCATCCCTTGTATCGGGAGTACTCTCTCT	1003
QY	929	GCATCATGTTTCATGGTGGCTTCCCTCCGTCTCGTCTCCACCATCTGATCTCAACTACCACC	988
Db	1004	TCATATGATTTTGTGCACCTTGTCATCGTTCATCAGTCTTTTGTGTCAACGTGCACT	1063
QY	989	ACCGGCAGCAGACACTCAGAAATGATGATTTGGATTGCTGTGGTGTTCCTTTATTGGC	1048
Db	1064	ACGAACTCCGACCACACACACGATGCCCACTTGGGTCAAGGCTGTGTGTTTGAACCTTC	1123
QY	1049	TGCCGTGGGTGCTCGCATGTCAAGGCG	1076
Db	1124	TCCCGAGGTCAATTTATGACTAGGCC	1151

RESULT 5

AK051730

LOCUS

DEFINITION

Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:DJ13068A06 Product:NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.

ACCSSION

AK051730
AK051730.1 GI:26342155

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)
Mus musculus

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Carninci, P. and Hayashizaki, Y.

ORGANISM

REFERENCE

AUTHORS

